

Document made available under the Patent Cooperation Treaty (PCT)

International application number: PCT/KR04/003517

International filing date: 30 December 2004 (30.12.2004)

Document type: Certified copy of priority document

Document details: Country/Office: KR
Number: 10-2004-0003610
Filing date: 17 January 2004 (17.01.2004)

Date of receipt at the International Bureau: 19 April 2005 (19.04.2005)

Remark: Priority document submitted or transmitted to the International Bureau in compliance with Rule 17.1(a) or (b)



World Intellectual Property Organization (WIPO) - Geneva, Switzerland
Organisation Mondiale de la Propriété Intellectuelle (OMPI) - Genève, Suisse



별첨 사본은 아래 출원의 원본과 동일함을 증명함.

This is to certify that the following application annexed hereto
is a true copy from the records of the Korean Intellectual
Property Office.

출원 번호 : 10-2004-0003610
Application Number

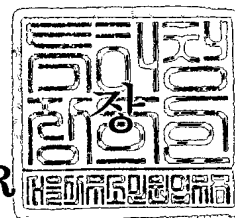
출원 년 월 일 : 2004년 01월 17일
Date of Application JAN 17, 2004

출원인 : 한국생명공학연구원
Applicant(s) Korea Research Institute of Bioscience and Biotechnology



2005 년 01 월 03 일

특 허 청
COMMISSIONER



【서지사항】

【서류명】	명세서 등 보정서
【수신처】	특허청장
【제출일자】	2004.01.26
【제출인】	
【명칭】	한국생명공학연구원
【출원인코드】	3-1999-034166-5
【사건과의 관계】	출원인
【대리인】	
【성명】	손민
【대리인코드】	9-1999-000420-6
【포괄위임등록번호】	2002-050197-2
【사건의 표시】	
【출원번호】	10-2004-0003610
【출원일자】	2004.01.17
【심사청구일자】	2004.01.17
【발명의 명칭】	재조합 단백질 생산용 맞춤형 단백질 융합인자의 초 고속 선별 방법
【제출원인】	
【접수번호】	1-1-2004-0020662-04
【접수일자】	2004.01.17
【보정할 서류】	명세서등
【보정할 사항】	
【보정대상항목】	별지와 같음
【보정방법】	별지와 같음
【보정내용】	별지와 같음
【취지】	특허법시행규칙 제13조·실용신안법시행규칙 제8조의 규 정에의하여 위와 같 이 제출합니다. 대리인 손민 (인)



1020040003610

출력 일자: 2005/1/4

【수수료】

【보정료】 0 원

【추가심사청구료】 0 원

【기타 수수료】 0 원

【합계】 0 원

【보정대상항목】 식별번호 97

【보정방법】 정정

【보정내용】

본 발명자들은 단백질을 합인자 TFP-1, TFP-2, TFP-3 및 TFP-4 유전자의 단편의 난발현성 단백질의 분비에 끼치는 영향을 조사하였으며, TFP-1의 경우 세린, 알라닌-리치 서열, N-글리코실화 부위 또는 이들 모두가 제거된 유전자를 포함하는 벡터의 경우 난발현성 단백질을 분비하지 못하였다. 반면, 5'-UTR(5'-untranslated region)을 제거하는 경우(pYIL-KRT1-4) 난발현성 단백질의 발현율이 3배 이상 증가하였으며, 추가의 3' 말단의 추가의 서열을 제거한 경우(pYIL-KRT1-3)에도 난발현성 단백질의 분비를 유도하였다.

【보정대상항목】 식별번호 98

【보정방법】 정정

【보정내용】

또 다른 양태로서, 본 발명은 서열번호 9로 기재되는 단백질을 합인자 TFP1-3 단백질 또는 이의 유사체에 관한 것이다. 또한, 본 발명은 단백질을 합인자 TFP1-3 단백질 또는 이의 유사체를 코딩하는 유전자에 관한 것이다. 달리, 본 발명은 서열번호 9로 기재되는 아미노산 서열 또는 이의 상동성, 바람직하게는 75%, 보다 바람직하게는

85%, 더욱 바람직하게는 90%, 가장 바람직하게는 95% 이상의 상동성을 갖는 아미노산 서열을 갖는 단백질융합인자 TFP1-3 단백질에 관한 것이다. 또한, 본 발명은 서열번호 9로 기재되는 단백질융합인자 TFP1-3 단백질을 코딩하는 DNA 서열 또는 이의 상동성, 바람직하게는 75%, 보다 바람직하게는 85%, 더욱 바람직하게는 90%, 가장 바람직하게는 95% 이상의 상동성을 갖는 DNA 서열을 갖는 유전자에 관한 것이다. 또한, 본 발명은 상기한 유전자를 포함하는 재조합 벡터에 관한 것이다. 바람직하게는, 재조합 벡터에 포함되는 유전자는 서열번호 9로 기재되는 단백질융합인자 TFP1-3을 코딩하는 유전자이다. 바람직하게는, 재조합 벡터는 pYIL-KRT1-3이다. 또한, 본 발명은 상기한 재조합 벡터로 형질전환된 세포에 관한 것이다. 바람직하게는, 형질전환된 세포는 pYIL-KRT1-3으로 형질전환된 에서리키아 콜라이 DH50/pYIL-KRT1-3(KCTC 10548BP)이다.

【보정대상항목】 식별번호 99

【보정방법】 정정

【보정내용】

또 다른 양태로서, 본 발명은 서열번호 10으로 기재되는 유전자에 의해 코딩되는 단백질융합인자 TFP1-4 단백질 또는 이의 유사체에 관한 것이다. 또한, 본 발명은 단백질융합인자 TFP1-4을 코딩하는 서열번호 10으로 기재되는 유전자 또는 이의 유사체에 관한 것이다. 달리, 본 발명은 서열번호 10으로 기재되는 유전자에 의해 코딩되는 아미노산 서열 또는 이의 상동성, 바람직하게는 75%, 보다 바람직하게는 85%, 더욱 바람직하게는 90%, 가장 바람직하게는 95% 이상의 상동성을 갖는 아미노산 서열을 갖는 단백질융합인자 TFP1-4 단백질에 관한 것이다. 또한, 본 발명은 서열번호 10로 기

재되는 DNA 서열 또는 이의 상동성, 바람직하게는 75%, 보다 바람직하게는 85%, 더욱 바람직하게는 90%, 가장 바람직하게는 95% 이상의 상동성을 갖는 DNA 서열을 갖는 유전자에 관한 것이다. 또한, 본 발명은 서열번호 10으로 기재되는 단백질을 합인자 TFP1-4을 코딩하는 유전자 또는 이의 유사체를 포함하는 재조합 벡터에 관한 것이다. 달리, 본 발명은 서열번호 10으로 기재되는 단백질을 합인자 TFP-1-4을 코딩하는 DNA 서열 또는 이의 상동성, 바람직하게는 75%, 보다 바람직하게는 85%, 더욱 바람직하게는 90%, 가장 바람직하게는 95% 이상의 상동성을 갖는 DNA 서열을 갖는 유전자를 포함하는 재조합 벡터에 관한 것이다. 바람직하게는, 재조합 벡터는 pYIL-KRT1-4이다. 또한, 본 발명은 상기한 재조합 벡터로 형질전환된 세포에 관한 것이다. 바람직하게는, 형질전환된 세포는 pYIL-KRT1-4으로 형질전환된 에서리키아 콜라이 DH5@/pYIL-KRT1-4(KCTC 10549BP)이다.

【보정대상항목】 식별번호 138

【보정방법】 정정

【보정내용】

상기 벡터 pYIL-KRT1-3 및 pYIL-KRT1-4는 2003년 11월 11일자로 국제기탁기관인 대전시 유성구 어은동 52번지 소재의 KCTC(Korean Collection for Type Cultures)에 각각 수탁번호 KCTC 10548BP 및 10549BP로 기탁하였다.

【서지사항】

【서류명】	특허출원서
【권리구분】	특허
【수신처】	특허청장
【제출일자】	2004.01.17
【발명의 명칭】	재조합 단백질 생산용 맞춤형 단백질융합인자의 초고속 선별 방법
【발명의 영문명칭】	Rapid screening method of suitable translational fusion partners for producing recombinant proteins
【출원인】	
【명칭】	한국생명공학연구원
【출원인코드】	3-1999-034166-5
【대리인】	
【성명】	손민
【대리인코드】	9-1999-000420-6
【포괄위임등록번호】	2002-050197-2
【발명자】	
【성명의 국문표기】	최의성
【성명의 영문표기】	CHOI, Eui Sung
【주민등록번호】	540923-1037424
【우편번호】	305-335
【주소】	대전광역시 유성구 궁동 395-3 다솔아파트 102동 507호
【국적】	KR
【발명자】	
【성명의 국문표기】	손정훈
【성명의 영문표기】	SOHN, Jung Hoon
【주민등록번호】	610728-1674115
【우편번호】	302-280
【주소】	대전광역시 서구 월평동 누리아파트 103동 506호
【국적】	KR
【발명자】	
【성명의 국문표기】	배정훈
【성명의 영문표기】	BAE, Jung Hoon
【주민등록번호】	660201-1386159



040003610

출력 일자: 2005/1/4

【우편번호】	427-010
【주소】	경기도 과천시 중앙동 주공아파트 1005동 503호
【국적】	KR
【발명자】	
【성명의 국문표기】	윤남경
【성명의 영문표기】	YUN,Nam Kyung
【주민등록번호】	781121-2448811
【우편번호】	306-130
【주소】	대전광역시 대덕구 평촌동 225-57호
【국적】	KR
【발명자】	
【성명의 국문표기】	성혜영
【성명의 영문표기】	SUNG,Hye Young
【주민등록번호】	791216-2408515
【우편번호】	301-132
【주소】	대전광역시 중구 문화2동 437-11
【국적】	KR
【심사청구】	청구
【미생물기탁】	
【기탁기관명】	KCTC
【수탁번호】	10544BP
【수탁일자】	2003.11.11
【미생물기탁】	
【기탁기관명】	KCTC
【수탁번호】	10545BP
【수탁일자】	2003.11.11
【미생물기탁】	
【기탁기관명】	KCTC
【수탁번호】	10546BP
【수탁일자】	2003.11.11
【미생물기탁】	
【기탁기관명】	KCTC
【수탁번호】	10547BP

【수탁일자】 2003.11.11

【미생물기탁】

【기탁기관명】 KCTC

【수탁번호】 10548BP

【수탁일자】 2003.11.11

【미생물기탁】

【기탁기관명】 KCTC

【수탁번호】 10549BP

【수탁일자】 2003.11.11

【핵산염기 및 아미노산 서열목록】

【서열개수】 39

【서열목록의 전자파일】 첨부

【취지】 특허법 제42조의 규정에 의한 출원, 특허법 제60조의 규정에 의한 출원심사를 청구합니다. 대리인 손민 (인)

【수수료】

【기본출원료】	69 면	38,000 원
【가산출원료】	0 면	0 원
【우선권주장료】	0 건	0 원
【심사청구료】	9 항	397,000 원
【합계】		435,000 원
【감면사유】	정부출연연구기관	
【감면후 수수료】	217,500 원	
【첨부서류】	1. 미생물기탁증명서[국제양식포함]_6통	

【요약서】

【요약】

본 발명은 재조합 생산이 어려운 난발현성 단백질(non-producible protein)을 발현 및 분비생산 할 수 있는 맞춤형 단백질융합인자(translational fusion partner: TFP)를 다양한 유전자원으로부터 초고속으로 선별하는 기술에 관한 것이다.

【대표도】

도 6

【색인어】

단백질융합인자(translational fusion partner: TFP), 효모, 난발현성 단백질(non-producible protein), 인버테이즈

【명세서】

【발명의 명칭】

재조합 단백질 생산용 맞춤형 단백질융합인자의 초고속 선별 방법{Rapid screening method of suitable translational fusion partners for producing recombinant proteins}

【도면의 간단한 설명】

도 1은 인버테이즈 유전자의 결실과정 및 선별표지의 팝-아웃 과정을 나타낸 그림이다.

도 2는 인버테이즈 활성측정을 위한 자이모그램이다.

레인 1,2,3 : 야생형 에스. 세레비지에(*S. cerevisiae*) 2805

레인 4,5,6 : 인버테이즈 결실변이주(*S. cerevisiae* 2805 Δ *inv2*)

도 3은 탄소원에 따른 균체의 성장을 보여주는 사진이다.

INV2 : 야생형 에스. 세레비지에 2805

Δ *inv2*: 인버테이즈 결실변이주(*S. cerevisiae* 2805 Δ *inv2*)

도 4는 인버테이즈 유전자 결실 확인을 위한 서던블롯팅 결과이다.

레인 1,2 : 에스. 세레비지에 Y2805 *ura3 INV2*,

레인 3,4 : 에스. 세레비지에 Y2805 Δ *inv2U* (*URA3* Δ *inv2*),

레인 5,6 : 에스. 세레비지에 Y2805 Δ *inv2* (*ura3* Δ *inv2*)

도 5는 글루코스 및 수크로즈 배지에서 균체성장 여부를 확인한 사진이다.

도 6은 플라스미드 pYHTS-F0, F1 및 F2의 제조과정 및 효모 유전자 라이브러리 제조과정을 도식한 그림이다.

도 7은 4종의 단백질융합인자를 포함하는 균체의 배양상등액을 SDS-PAGE 및 웨스턴브롯팅한 결과이다.

레인 1: 크기 표시자(size marker)

레인 2: 인터류킨-2

레인 3: pYIL-TFP1를 함유한 균체를 배양한 배양상등액

레인 4: pYIL-TFP2를 함유한 균체를 배양한 배양상등액

레인 5: pYIL-TFP3를 함유한 균체를 배양한 배양상등액

레인 6: pYIL-TFP4를 함유한 균체를 배양한 배양상등액

도 8은 당쇄분석을 위해 엔도-에이치(Endo-H) 처리 전후의 SDS-PAGE 결과이다.

레인 1, -: pYIL-TFP1를 함유한 균체를 배양한 배양상등액, 엔도에이치 비처리

레인 1, +: pYIL-TFP1를 함유한 균체를 배양한 배양상등액, 엔도에이치 처리

레인 2, -: pYIL-TFP3를 함유한 균체를 배양한 배양상등액, 엔도에이치 비처리

레인 2, +: pYIL-TFP3를 함유한 균체를 배양한 배양상등액, 엔도에이치 처리

레인 3, -: pYIL-TFP4를 함유한 균체를 배양한 배양상등액, 엔도에이치 비처리

레인 3, +: pYIL-TFP4를 함유한 균체를 배양한 배양상등액, 엔도에이치 처리

도 9는 Kex2p 프로세싱 사이트 유무에 따른 균체 배양상등액의 SDS-PAGE 분석 결과이다.

레인 M: 크기 표시자

레인 1: pYIL-TFP1를 함유한 균체의 배양상등액

레인 2: pYIL-KRTFP1를 함유한 균체의 배양상등액

레인 3: pYIL-TFP3를 함유한 균체의 배양상등액

레인 4: pYIL-KRTFP3를 함유한 균체의 배양상등액

레인 5: pYIL-TFP4를 함유한 균체의 배양상등액

레인 6: pYIL-KRTFP4를 함유한 균체의 배양상등액

도 10은 TFP-1의 특성분석을 위한 유전자 결실 후 제조된 플라스미드들을 나타낸 모식도이다.

도 11은 TFP-1으로부터 유래된 단백질융합인자들(TFP1-1, 2, 3 및 4)에 의한 인터류킨-2 분비능을 확인한 SDS-PAGE 결과이다.

레인 M: 크기 표시자

레인 S: 인터류킨-2

레인 1-1: pYIL-KRT1-1을 함유한 균체를 배양한 배양상등액

레인 1-2: pYIL-KRT1-2를 함유한 균체를 배양한 배양상등액

레인 1-3: pYIL-KRT1-3을 함유한 균체를 배양한 배양상등액

레인 1: pYIL-KRTFP1을 함유한 균체를 배양한 배양상등액

레인 1-4: pYIL-KRT1-4를 함유한 균체를 배양한 배양상등액

도 12는 단백질융합인자 TFP-1을 이용한 난분비 단백질 인체 콜로니자극인자(G-CSF)의 분비를 확인한 SDS-PAGE 및 웨스턴브롯팅 결과이다.

레인 M: 크기 표시자

레인 1: pYIL-KRTFP1을 함유한 균체를 배양한 배양상등액

레인 2: pYGCSF-TFP1을 함유한 균체를 배양한 배양상등액

레인 3: pYGCSF-MF α 를 함유한 균체를 배양한 배양상등액

【발명의 상세한 설명】

【발명의 목적】

【발명이 속하는 기술분야 및 그 분야의 종래기술】

- <50> 본 발명은 효모에서 재조합 생산이 어려운 단백질을 발현 및 분비생산 할 수 있는 맞춤형 단백질융합인자(translational fusion partner: TFP)를 다양한 유전자원으로부터 초고속으로 선별하는 방법에 관한 것이다.
- <51> 최근 인체게놈 프로젝트에서 확보된 유전체 염기서열 정보와 유전체 단위에서 밝혀지는 다양한 단백질의 기능을 분석하고 인체 의약학적으로 중요한 단백질 제품생산을 위해서는 재조합 미생물을 이용하는 고효율 단백질 생산 시스템 개발이 필요하다. 인체와 같은 고등생물 유래의 재조합단백질을 생산하기 위해서 발현시스템을 선정할 때 숙주세포의 성장특성, 단백질 발현정도, 세포내외 발현가능성, 번역 후 수식(post-translational modification) 가능성, 발현된 단백질의 생물학적 활성 및 발현단백질의 용도 등과 같은 다양한 요인들이 고려되어야 한다. 대표적 미생물 유전자 발현시스템으로 에شري키아 콜라이 및 효모 시스템이 주로 이용되고 있는데 에شري키아 콜라이는 많은 발현시스템이 개발되어 있고 외래단백질의 발현율이 매우 높은 장점이 있지만 고등생물 유래의 단백질을 재조합 생산하고자 할 때 번역 후 수식 과정이 불가능하며 세포의 배양배지로 단백질의 완전한 분비가 어렵고 이황화 결합(disulfide bond)이 많은 단백질의 폴딩(folding)이 불가능하며 봉합체(inclusion body) 등의 불용성 단백질 형태로 생산하는 등의 단점이 지적되고 있다 (Makrides, Microbial Rev., 1996, 60, 512). 또한,

인체단백질 중 질병과 연관되어 의학학적으로 가치가 높은 대부분의 단백질이 당단백질이거나 막단백질이기 때문에 완전한 활성을 갖기 위해서는 글리코실화(glycosylation)를 반드시 요구하거나 이황화 결합을 통한 완전한 3차원 구조를 요구할 경우 에شري키아 콜라이에서는 생산이 불가능하며 효모와 같은 진핵미생물 발현 시스템을 반드시 필요로 한다.

<52> 진핵 미생물인 효모 사카로마이세스 세레비시아(*Saccharomyces cerevisiae*)는 인체에 대해 안전성이 입증된 GRAS(Generally Recognized As Safe) 미생물로서 유전자 조작이 용이하며 다양한 발현시스템이 개발되어 있고 대량배양이 용이하다. 뿐만 아니라 인체단백질과 같은 고등세포 유래의 단백질을 재조합 생산할 때 단백질을 세포 밖으로 분비할 수 있는 분비기능과 글리코실화 등과 같은 단백질의 번역 후 수식 기능을 수행할 수 있는 장점을 제공한다. 단백질의 분비시그널과 목표단백질을 인위적으로 융합(fusion)함으로써 세포의 분비가 가능한데 단백질의 분비과정을 통해서 단백질의 폴딩이나 이황화 결합의 형성 및 당쇄부가 과정이 진행되며 따라서 생물학적으로 완전한 활성을 갖는 재조합단백질을 생산할 수 있는 장점을 제공한다. 이는 또한 생물학적 활성을 갖는 단백질을 배지로부터 직접 얻을 수 있기 때문에 경제적으로 효율이 낮은 세포의 분쇄나 재접힘 단계를 필요로 하지 않아 매우 경제적이다 (Eckart and Bussineau, Curr. Opin. Biotechnol., 1996, 7, 525).

<53> 그러나 상기한 많은 장점에도 불구하고 효모 사카로마이세스 세레비시아를 이용한 인체 단백질 분비시스템에 관한 현 기술의 문제점으로 지적되고 있는 것이 인체단백질의 종류에 따라서 전혀 생산되지 않거나 수 그램/리터까지 생산되는 등 분비율이 수천 배 이상 차이를 보여 분비생산성을 예측하기 힘든 것이다. 외래단백질이 수 그램/리터 수준까지 분비 생산이 가능한 것으로 판단할 때 분비 생산성 측면에서 충분한 경제성이 있는 것으로 판단되나 단백질의 종류에 따라서 분비효율이 낮은 문제와 특히 고부가가치의 인체의약품 단백질을 생산하고자 할

때 발현 및 분비가 힘든 문제가 자주 발생한다. 이러한 문제를 해결하기 위해서 단백질의 분비에 관여하는 분비인자에 대한 연구가 많이 진행되고 있다. 소포체에서 새로 합성된 단백질의 폴딩을 돕는 분비인자인 BiP(*KAR2*)의 과발현 방법(Robinson 등, *Biotechnol. prog.*, 1996, 271, 10017)과 시스테인 결합의 형성을 돕는 *PDI*(protein disulfide isomerase)의 과발현 방법(Robinson 등, *Bio/Technology*, 1994, 12, 381; Schultz 등, *Ann. N. Y. Acad. Sci.*, 1994, 721, 148; Hayano 등, *FEBS Lett.*, 1995, 377, 505) 등의 chaperone에 대한 연구 및 분비를 유도하는 융합인자의 제조와 원래 분비가 잘되는 단백질들과의 융합을 통해서 분비를 증진시키는 연구가 많이 진행되고 있다(Gouka 등, *Appl. Microbiol. Biotechnol.*, 1997, 47, 1). 이러한 방법은 현재까지 외래단백질의 분비를 증진시킨 매우 성공적인 방법으로 인식되고 있는데 이러한 융합기술의 분자 기작에 대한 연구는 아직 미비하지만 실험적으로 이러한 융합기술은 단백질의 이동을 용이하게 하고 폴딩을 돕는 등 단백질의 번역단계나 번역후 단계에서의 한계점을 개선하는 것으로 알려져 있다.

<54> Kjeldsen 등 (*Protein Expr. Purif.*, 1997, 9, 331)은 인슐린 또는 인슐린 전구체 (insulin precursor, IP)를 효모에서 효율적으로 분비 생산하기 위하여 이론적 근거에 의해 제작된 합성 리더(leader)를 인슐린에 융합하여 인슐린의 분비율을 개선하였는데 이러한 합성 리더에는 글리코실화 부위(N-glycosylation site)와 BiP 인식부위를 첨가하여 소포체에서 머무르는 시간을 길게 함으로써 단백질이 제대로 폴딩이 되도록 유도하였다. 또한 합성 리더에 추가적인 당쇄부가 부위를 도입하여 아스퍼질러스 나이저 및 사카로마이세스 세레비지애에서 인슐린의 분비율이 상당히 증가되었다고 보고하였다(Kjeldsen 등, *Protein Expr. Purif.*, 1998, 14, 309). 유사한 결과가 아스퍼질러스 아와모리(Ward 등, *Bio/Technology*, 1989, 8, 435) 및 소수성 큐티나제(cutinase)를 효모에서 발현한 경우에도 보고되었다(Sagt 등, *Appl. Environ.*

Microbiol. 2000, 66, 4940). 이는 글리코실화 부위의 도입으로 인하여 재조합 단백질의 소포체내에서의 용해성이 증가되고 단백질의 폴딩이 유도되어 결과적으로 분비가 증가되는데 기인하였다.

<55> 원래 분비가 잘되는 단백질을 융합파트너로 이용하는 연구로는 아스퍼질러스 나이지의 글루코아밀레이즈와 융합발현하여 소 유래의 프로카이모신(bovine prochymosin) (Ward 등, Bio/Technology, 1989, 8, 435), 돼지 췌장유래의 포스포리파제 A2(porcine pancreatic phospholipase A2) (Roberts 등, Gene, 1992, 122, 155), 인체 인터류킨-6(human interleukin-6) (Contreras 등, Bio/Technology 1991, 9, 378; Broekhuijsen 등, J. Biotechnol., 1993, 31, 135), 닭 유래의 라이소자임(hen egg-white lysozyme) (Jeenes 등, FEMS Microbiol Lett, 1993, 107, 267) 및 인체 락토페린(human lactoferrin) (Ward 등, Bio/Technology, 1995, 13, 498)을 효율적으로 분비시켰다. 분비증가율은 단백질에 따라 차이를 보여 5배에서 1000배까지 증가하였다. 효모에서도 인체 인터류킨-1베타(interleukin-1 β)의 아미노말단 24개의 아미노산을 융합파트너로 이용하여 인체 성장인자(human growth hormone) 및 콜로니 자극인자(Granulocyte colony-stimulating factor, G-CSF)의 분비를 증가시켰다 (Lee 등, Biotechnol. Prog., 1999, 15, 884). 인체 인터류킨-1베타는 특별한 분비 시그날이 없이 분비되는 것으로 알려져 있으며 (Muesch 등, Trends Biochem. Sci., 1990, 15, 86) 효모에서도 매우 효과적으로 재조합 분비생산되는 것으로 보고되었다 (Baldari 등, Protein Eng., 1987, 1, 433). 단백질이 제대로 폴딩하기 위해서는 단백질이 원래 보유한 융합 파트너가 반드시 필요한 경우도 보고되었는데 (Takahashi 등, Appl Microbiol. Biotechnol., 2001, 55, 454) 라이조퍼스 오라이제 유래의 리파제(ROL)를 효모 사카로마이세스

세레비시에에서 발현하기 위해 사카로마이세스 유래의 교배인자 알파(mating factor alpha)의 프리-프로-리더(pre-pro-leader) 서열을 ROL의 성숙단백질에 해당하는 유전자와 결합하여 발현한 경우에는 전혀 ROL이 분비되지 않았으나 ROL 자체의 pro 서열을 결합한 경우에는 적절하게 분비되었다. 이는 ROL 자체의 pro 서열이 자기 단백질의 폴딩에 절대적으로 필요하다는 것을 보여주었다.

<56> 상기한 연구결과에서 보는 바와 같이 재조합단백질의 분비를 유도하기 위해서 다양한 분비인자가 개발되었다. 그러나 개발된 분비인자들이 특정단백질의 분비증진에는 효과가 있으나 모든 단백질에 일률적으로 적용될 수 없는 문제가 있었다. Dorner 등은 CHO 세포에서 BiP를 과발현한 결과 오히려 단백질 분비율이 감소하였다고 보고하였고 (Dorner 등, EMBO J., 1992, 11, 1563), 반면 BiP의 발현을 감소시켰을 때 단백질 분비가 증가하였다 (Dorner 등, Mol. cell. Biol., 1988, 8, 4063). 효모에서도 *KAR2*(BiP)의 과발현이 식물 타우마틴(plant thaumatin)의 경우에는 분비가 향상되지 않았다 (Harmsen 등, Appl. Microbiol. Biotechnol., 1996, 46, 365). 배큘로바이러스에서 BiP를 과발현하였을 때 세포추출물에서 가용성 항체의 양이 증가하였으나 항체의 분비효율은 증가하지 않았다 (Hsu 등, Protein Expr. Purif., 1994, 5, 595). 또 다른 분비인자인 폴데이즈(*PDI*)를 과발현하는 경우에도, 아스퍼질러스 나이지에서 폴데이즈를 과발현 하였으나 글루코아밀레이즈의 분비가 증가하지 않았다 (Wang and Ward, Curr. Genet. 2000, 37, 57). 또한 단백질 융합파트너를 이용한 분비유도의 경우에도 특정 단백질에만 분비효율이 증진되는 문제가 보고되고 있다.

【발명이 이루고자 하는 기술적 과제】

<57> 상기한 결과에서 보는 바와 같이 분비인자의 효과에 대한 많은 연구가 있었으나 단백질의 종류에 따라서 분비에 미치는 영향이 서로 상이했기 때문에 개발된 분비인자가 모든 단백질에 적용될 수 없는 문제가 있다. 따라서 목적하는 각 단백질 분비증진 최대화를 위해서는 목적단백질에 특이적으로 적용될 수 있는 최적의 분비인자를 선별하는 기술이 필요하다. 이에 본 발명자들은 재조합단백질의 종류에 따라서 최적의 분비융합인자를 유전체 단위에서 초고속으로 선별하는 기술을 개발하여 본 발명을 완성하였다.

<58> 따라서, 본 발명의 목적은 효모에서 발현율이 낮아 경제적 대량 생산이 불가능한 단백질을 대상으로 단백질 생산을 강력하게 유도할 수 있는 적절한 단백질융합인자(TFP)를 효모 유전체 및 다양한 유전자원으로부터 초고속 발굴할 수 있는 방법의 제공과 이를 이용하여 난발현성인 단백질의 분비생산을 촉진할 수 있는 방법을 제공하는 것이다.

【발명의 구성 및 작용】

<59> 하나의 양태로서, 본 발명은 난발현성 목적 단백질 유전자(X)를 자동선별용 리포터 유전자(R)에 결합하여 X-R의 융합물을 제조하고, X-R의 융합물에 또 다른 유전자 융합을 통해 X-R 융합 단백질을 세포 밖으로 분비를 유도하는 단백질융합인자(TFP)를 유전자 라이브러리로부터 선별하는 방법에 관한 것이다.

<60> 바람직하게는, 본 발명은

<61> (1) 난발현성 목적 단백질의 유전자(X)와 인프레임(in-frame)으로 연결된 자동선별용 리포터 유전자(R)의 융합 유전자(X-R)를 포함하는 자동 선별 벡터를 제조하는 단계;

- <62> (2) 난발현성 융합 단백질(X-R)의 분비를 유도하는 단백질융합인자(translational fusion partner: TFP)를 포함한 게놈성 DNA 또는 cDNA를 자동 선별 벡터와 연결하여 단백질융합인자 라이브러리를 제조하는 단계;
- <63> (3) 단백질융합인자 라이브러리로 리포터 유전자의 활성이 없는 세포를 형질전환시켜 리포터 단백질의 활성을 검출하는 단계; 및
- <64> (4) 리포터 단백질의 활성을 나타내는 형질전환 세포로부터 유전자를 분리하여 단백질융합인자의 특성을 분석하는 단계를 포함하여, 난발현성 단백질 생산용 맞춤형 단백질융합인자를 선별하는 방법에 관한 것이다.
- <65> 보다 바람직하게는, 본 발명은
- <66> (1) 효모 인버테이즈를 이용한 자동선별 시스템 개발을 위한 리포터 유전자로 사용하기 위하여 효모 자체가 갖는 인버테이즈 유전자인 INV2(I)를 결실시킨 효모변이주를 제조하는 단계;
- <67> (2) 효모 *GAL10* 프로모터에 의해서 발현이 조절되며 인버테이즈(I) 유전자와 난발현성 유전자(X)가 인프레임(in-frame)으로 융합된 유전자(X-I)를 함유하는 자동선별 벡터인 효모 HTS (high throughput selection) 벡터 (pYHTS-F0, pYHTS-F1 및 pYHTS-F2)의 제조 단계;
- <68> (3) 인버테이즈와 난발현성 단백질의 융합 유전자(X-I)를 분비시킬 수 있는 효모 유전체로부터 단백질융합인자 라이브러리를 pYHTS 벡터에 제조하는 단계;
- <69> (4) 제조된 라이브러리를 단계 (1)에서 제조된 효모에 형질전환하고 수크로스를 단일탄소원으로 포함하고 있는 배지에서 자동선별하는 단계;

- <70> (5) 수크로즈 배지에서 성장한 효모를 배양하여 배지로 분비된 단백질을 확인하는 단계;
및
- <71> (6) 효모로부터 유전자를 분리하여 단백질융합인자 특성을 분석하는 단계를 포함하는 난발현성 단백질 생산용 맞춤형 TFP를 선별하는 방법에 관한 것이다.
- <72> 제조합 생산이 힘든 난발현성 단백질(X)와 인버테이즈(I)를 융합하여 효모에서 발현할 때 융합된 난발현성단백질(X)로 인하여 정상적으로 분비되던 인버테이즈(I)의 분비가 억제되고 따라서 수크로즈를 단일탄소원으로 함유하는 배지에서 난발현의 정도에 따라서 효모가 성장하지 못하거나 또는 성장이 매우 지연된다. 그러나, X-I의 발현 및 분비를 유도할 수 있는 효율적인 단백질융합인자를 도입할 경우 세포는 수크로즈 배지에서 빠르게 성장한다. 이러한 특징을 이용하여 난발현단백질(X)과 인버테이즈(I)가 융합된 단백질(X-I)에 다양한 유전체로부터 확보되는 단백질융합인자 라이브러리를 추가적으로 융합하여 TFP-X-I 또는 X-I-TFP의 형태로 제조한 후 효모에 도입, 발현하여 수크로즈 배지에서 빠르게 성장하는 세포를 선별하면 다양한 유전체 라이브러리로부터 난발현단백질에 가장 적합한 TFP 를 초고속으로 선별할 수 있다.
- <73> 구체적으로, 본 발명자는 효모 인버테이즈 결여 변이주 제조하고, 인버테이즈 유전자가 결실된 효모 균주에서 인버테이즈와 융합된 단백질의 발현을 통해 인버테이즈를 자동선별 시스템의 마커로 사용할 수 있음을 확인하였다. 이어, 난발현성 단백질인 인체 인터류킨-2를 이용한 단백질융합인자 자동선별 벡터인 pYHTS-F0, F1 및 F2를 제조하고, 이에 효모 기원의 절단된 염색체 DNA를 연결하여 단백질융합인자 라이브러리 제조하고, 이로부터 난분비성 단백질 인체 인터류킨-2에 적합한 단백질융합인자 단백질 TFP-1, TFP-2, TFP-3 및 TFP-4를 확인하였다.
- <74> 본 발명에서 "단백질융합인자(translational fusion partner: TFP)"란 난발현성 단백질을 코딩하는 유전자와 융합되어 난발현성 단백질의 분비생산을 유도하는 유전자를 의미한다. 또

한, "단백질융합인자 단백질"은 상기한 바와 단백질융합인자 유전자에 의해 코딩되는 아미노산 서열의 단백질을 의미한다. 구체적으로, TFP-1, TFP-2, TFP-3, TFP-4 등을 열거할 수 있다.

<75> 본 발명에서 "인버테이즈를 이용한 자동선별 시스템"이란 효모가 수크로스를 단일탄소원으로 이용하기 위해서는 효모 *INV2* 유전자에 의해서 코딩되는 단백질인 인버테이즈라는 효소를 필요로 하는데 효모의 염색체 상에 존재하는 *INV2* 유전자를 결실시키고 벡터에 *INV2* 유전자를 도입하여 *INV2* 유전자의 발현 여부에 따라서 수크로스 배지에서 성장하는 균주를 선별하는 시스템을 의미한다.

<76> 본 발명의 자동선별 리포터 유전자는, 반드시 이에 한정되는 것은 아니나, 인버테이즈, 수크레이즈, 셀룰레이즈, 자일라네이즈, 말테이즈, 아밀레이즈, 글루코아밀레이즈, 갈락토시데이즈 등으로 구성된 군으로부터 선택된다.

<77> 본 발명에서 "난발현성 단백질"이란 인체 또는 다양한 생명체 유래의 단백질을 재조합 생산하고자 할 때 단백질 자체의 특성으로 인해서 에شري키아 콜라이이나 효모 등의 숙주세포에서 재조합 발현 생산하기 어려운 단백질을 의미한다. 본 발명의 목적상, 난발현성 단백질은 바람직하게는 에شري키아 콜라이 등 다른 숙주에서 재조합 생산이 가능하더라도 효모에서는 생산성이 낮아 경제성이 없는 다수의 단백질을 포함할 수 있다.

<78> 본 발명에서 확보된 단백질융합인자 TFP-1, TFP-2, TFP-3, TFP-4, TFP1-3 및 TFP1-4의 적용범위는 인체 인터류킨-2(IL-2) 및 인체 콜로니자극인자(G-CSF)를 포함하여 다양한 상업적 용도로 대량 생산되는 단백질을 포함한다. 이러한 단백질들은, 이들로 한정되는 것은 아니지만, 혈청단백질(예, 인자 VII, VIII 및 IX를 포함한 혈액인자), 면역글로불린, 사이토카인(예, 인터류킨), α -, β - 및 γ -인터페론, 콜로니자극인자(GM-CSF), 혈소판 유도된 성장 인자(PDGF), 포스포리파제-활성화 단백질(PLAP), 인슐린, 종양 괴사 인자(TNF), 성장 인자(예, TGF

α 또는 $TGF\beta$ 와 같은 조직 성장 인자 및 내피 성장 인자), 호르몬(예, 소낭-자극 호르몬, 갑상선-자극 호르몬, 항이노 호르몬, 색소성 호르몬 및 부갑상선 호르몬, 황체호르몬 분비 호르몬 및 이의 유사체), 칼시토닌(calcitonin), 칼시토닌 유전자 관련 펩타이드(Calcitonin Gene Related Peptide; CGPR), 엔케팔린(enkephalin), 소마토메딘, 에리스로포이에틴, 시상하부 분비 인자, 프롤락틴, 만성 고나도트로핀, 조직 플라스미노겐 활성화제, 성장호르몬 분비 펩타이드(growth hormone releasing peptide; GHPR), 흉선 체액성 인자(thymic humoral factor; THF) 등이 포함된다. 또한, 이러한 단백질에는 효소를 포함할 것이며, 예로는 탄수화물-특이적 효소, 단백질분해 효소, 산화환원 효소, 트랜스퍼라제, 하이드롤라제, 라이아제, 이소머라제 및 리가제가 포함된다. 구체적인 효소로는 이들로 한정하는 것은 아니지만 아스파라기나제, 아르기나제, 아르기닌 데아미나제, 아데노신 데아미나제, 과산화물 디스뮤타제, 엔도톡시나제, 카탈라제, 키모트립신, 리파제, 우리카제, 아데노신 디포스파타제, 티로시나제 및 빌리루빈 옥시다제를 들 수 있다. 탄수화물-특이적 효소의 예로는 글루코즈 옥시다제, 글루코다제, 갈락토시다제, 글루코세레브로시다제, 글루코우로니다제 등이 포함된다.

<79> 난발현성 단백질 유전자는 인체 의학적 또는 산업적 중요성이 있으며 재조합 생산의 필요성이 있는 인체를 포함한 다양한 동식물 및 미생물 유래의 유전체 또는 cDNA로부터 선택되거나 화학 합성되는 되는 것으로서, 상기한 단백질들을 코딩하는 유전자이다.

<80> 본 발명의 자동 선별 벡터는 프로모터 유전자, 번역개시 및 종결코돈이 제거된 목적 단백질을 코딩하는 유전자 및 이에 인프레임으로 융합된 리포터 유전자를 포함하며, 프로모터 유전자는

GAPDH, *PGK*, *ADH*, *PHO5*, *GAL1* 및 *GAL10*으로 구성된 군으로부터 선택되는 유전자인 것이 바람직하며 난발현성 단백질에 적합한 맞춤형 융합인자 발굴을 위한 융합인자 라이브러리는 효모 또는 인체를 포함한 다양한 동식물 및 미생물 유래의 유전체 또는 cDNA로부터 선택되거나 화학 합성되는 것을 포함하며, 바람직하게는 효모류 유전자로부터 기원한다.

<81> 본 발명의 자동선별 방법에서 형질전환에 사용한 숙주세포는 캔디다(*Candida*), 디베리오 마이세스(*Debaryomyces*), 한세놀라(*Hansenula*), 클루이베로마이세스(*Kluyveromyces*), 피키아(*Pichia*), 스키향사카로마이세스(*Schizosaccharomyces*), 야로이아(*Yarrowia*) 및 사카로마이세스(*Saccharomyces*) 속 등의 효모류, 아스퍼질러스(*Aspergillus*), 페니실리엄(*Penicillium*), 라이조푸스(*Rhizopus*) 및 트리코더마(*Trichoderma*) 속 등의 균류 또는 에셔리키아(*Escherichia*) 및 바실러스(*Bacillus*) 속 등의 세균류를 사용할 수 있으나 반드시 이에 한정되는 것은 아니다.

<82> 본 발명의 난발현성 단백질 생산용 맞춤형 TFP의 초고속 선별 방법은 발현이 불가능하거나 발현율이 매우 낮은 난발현성 단백질용으로 사용하는 것이 적합하지만 발현이 어느 정도 되지만 발현율을 높일 수 있는 TFP를 선별하기 위해서도 이용될 수 있다. 본 발명의 구체예에서 실시하고 있는 바와 같이 인버테이즈를 리포터로 사용할 경우 수크로스 배지에서 성장이 빠른 순서대로 구별하여 보다 효율적인 TFP를 선별할 수 있다.

<83> 또 다른 양태로서, 본 발명은 난발현성 단백질 인터류킨-2의 분비생산 촉진을 위한 맞춤형 융합인자 초고속 선별백터 pYHTS-F0, F1 및 F2에 관한 것이며, 이러한 선별 백터는 난발현 단백질 인체 인터류킨-2와 인버테이즈가 융합된 유전자를 함유하고 있으며 인터류킨-2 유전자의 아미노말단에 세 가지 서로 다른 리딩프레임으로 제조된 제한효소 *Bam*HI 절단부위를 함유하고 있다.

<84> 본 발명의 구체적 실시에서는 인체 인터류킨-2를 효모에서 분비생산 촉진하는 맞춤형 융합인자를 초고속 선별하기 위하여 선별벡터 3 종(pYHTS-F0, F1 및 F2)에 효모 염색체를 무작위로 절단한 후 삽입한 다음 인버테이즈가 결여된 균주에 형질전환하고 수크로스 배지에서 성장하는 균체를 선별함으로써 난발현성 인터류킨-2 및 인버테이즈의 융합단백질을 배지로 분비할 수 있는 맞춤형 융합인자를 선별하였다.

<85> 인체 인터류킨-2는 소수성이 강한 단백질로서 효모에서 발현이 어려운 이유는 강력한 프로모터에 의해서 재조합 대량 발현된 단백질이 소포체내에서 빠르게 활성형으로 폴딩(folding)되지 못하고 서로 응집되어 소포체 기능을 마비시키는 문제 때문으로 추정되고 있다. 따라서 인터류킨-2에 융합된 인버테이즈도 분비되지 못해 세포가 수크로스 배지에서 성장할 수 없다. 이러한 융합단백질을 효율적으로 분비시킬 수 있는 단백질융합인자는 인터류킨-2 유전자의 앞부분에 효모 유전체 라이브러리를 삽입하고 효모에 형질전환한 후 수크로스 배지에서 성장하는 형질전환체를 확보함으로써 가능하게 된다.

<86> 본 발명을 이용하여 난발현단백질인 인터류킨-2의 분비를 유도하는 융합인자를 확보하기 위하여 수크로스 배지에서 성장하는 형질전환체로부터 유전자를 분리하고 에서리키아 콜라이으로 재형질전환하여 서로 다른 네 종류의 플라스미드(pYHTS-TFP1, TFP2, TFP3 및 TFP4)를 회수하였다. 각각의 플라스미드에 삽입된 4종의 서로 다른 단백질융합인자 유전자 TFP-1(서열번호 2), TFP-2(서열번호 4), TFP-3(서열번호 6) 및 TFP-4(서열번호 8)를 확보하였으며, 이들 유전자에 의해 코딩되는 아미노산 서열이 각각 서열번호 1, 3, 5 및 7에 나타나 있다.

<87> 상기 수득된 벡터 pYHTS-TFP1, TFP2, TFP3 및 TFP4로부터 인버테이즈를 제거하고 인터류킨-2 유전자에 번역종결코돈을 삽입한 pYIL-TFP1, TFP2, TFP3 및 TFP4를 제작하였으며, 이러한 벡터는 단백질융합인자와 융합된 형태의 인터류킨-2를 분비하기 때문에 단백질융합인자를 자

동제거할 수 있도록 단백분해효소 Kex2p 인식부위를 삽입한 벡터 pYIL-KRTFP1, KRTFP-2, KRTFP-3 및 KRTFP-4를 제작하였다. 나아가, 인체 콜로니 자극인자(G-CSF)를 단백질융합인자 TFP-1와 융합한 벡터 pYGCSF-TFP1을 제작하여 TFP-1이 인체 인터류킨-2 이외의 단백질의 분비 생산에도 효과적임을 입증하였다.

<88> 따라서, 또 다른 양태로서, 본 발명은 서열번호 1로 기재되는 단백질융합인자 TFP-1 단백질, 이의 유사체 또는 이의 단편에 관한 것이다. 또한, 본 발명은 서열번호 1로 기재되는 단백질융합인자 TFP-1 단백질, 이의 유사체 또는 이의 단편을 코딩하는 유전자에 관한 것이다. 달리, 본 발명은 서열번호 1로 기재되는 아미노산 서열 또는 이의 상동성(homologous), 바람직하게는 75%, 보다 바람직하게는 85%, 더욱 바람직하게는 90%, 가장 바람직하게는 95% 이상의 상동성을 갖는 아미노산 서열을 갖는 단백질융합인자 TFP-1 단백질에 관한 것이다. 또한, 본 발명은 서열번호 1로 기재되는 단백질융합인자 TFP-1 단백질을 코딩하는 DNA 또는 이의 상동성, 바람직하게는 75%, 보다 바람직하게는 85%, 더욱 바람직하게는 90%, 가장 바람직하게는 95% 이상의 상동성을 갖는 DNA 서열을 갖는 유전자에 관한 것이다. 바람직하게는, 상기한 유전자는 서열번호 2의 유전자이다. 또한, 본 발명은 상기한 유전자를 포함하는 재조합 벡터에 관한 것이다. 바람직하게는, 재조합 벡터에 포함되는 유전자는 서열번호 2의 유전자이다. 바람직하게는, 재조합 벡터는 pYIL-TFP1, pYIL-KRTFP1 또는 pYGCSF-TFP1이다. 또한, 본 발명은 상기한 재조합 벡터로 형질전환된 세포에 관한 것이다. 바람직하게는, 형질전환된 세포는 pYIL-KRTFP1으로 형질전환된 에شري키아 콜라이 DH50/pYIL-KRTFP1(KCTC 10544BP) 이다.

<89> 또 다른 양태로서, 본 발명은 서열번호 3으로 기재되는 단백질융합인자 TFP-2 단백질, 이의 유사체 또는 이의 단편에 관한 것이다. 또한, 본 발명은 서열번호 3으로 기재되는 단백질융합인자 TFP-2 단백질, 이의 유사체 또는 이의 단편을 코딩하는 유전자에 관한 것이다. 달

리, 본 발명은 서열번호 3로 기재되는 아미노산 서열 또는 이의 상동성, 바람직하게는 75%, 보다 바람직하게는 85%, 더욱 바람직하게는 90%, 가장 바람직하게는 95% 이상의 상동성을 갖는 아미노산 서열을 갖는 단백질융합인자 TFP-2 단백질에 관한 것이다. 또한, 본 발명은 서열번호 3으로 기재되는 단백질융합인자 TFP-2 단백질을 코딩하는 DNA 서열 또는 이의 상동성, 바람직하게는 75%, 보다 바람직하게는 85%, 더욱 바람직하게는 90%, 가장 바람직하게는 95% 이상의 상동성을 갖는 DNA 서열을 갖는 유전자에 관한 것이다. 바람직하게는, 상기한 유전자는 서열번호 4의 유전자이다. 또한, 본 발명은 상기한 유전자를 포함하는 재조합 벡터에 관한 것이다. 바람직하게는, 재조합 벡터에 포함되는 유전자는 서열번호 4의 유전자이다. 바람직하게는, 재조합 벡터는 pYIL-TFP2 또는 pYIL-KRTFP2이다. 또한, 본 발명은 상기한 재조합 벡터로 형질전환된 세포에 관한 것이다. 바람직하게는, 형질전환된 세포는 pYIL-KRTFP2로 형질전환된 에شري키아 콜라이 DH5@/pYIL-KRTFP2(KCTC 10545BP)이다.

<90> 또 다른 양태로서, 본 발명은 서열번호 5로 기재되는 단백질융합인자 TFP-3 단백질, 이의 유사체 또는 이의 단편에 관한 것이다. 또한, 본 발명은 서열번호 5로 기재되는 단백질융합인자 TFP-3 단백질, 이의 유사체 또는 이의 단편을 코딩하는 유전자에 관한 것이다. 달리, 본 발명은 서열번호 5로 기재되는 아미노산 서열 또는 이의 상동성, 바람직하게는 75%, 보다 바람직하게는 85%, 더욱 바람직하게는 90%, 가장 바람직하게는 95% 이상의 상동성을 갖는 아미노산 서열을 갖는 단백질융합인자 TFP-3 단백질에 관한 것이다. 또한, 본 발명은 서열번호 5로 기재되는 단백질융합인자 TFP-3 단백질을 코딩하는 DNA 서열 또는 이의 상동성, 바람직하게는 75%, 보다 바람직하게는 85%, 더욱 바람직하게는 90%, 가장 바람직하게는 95% 이상의 상동성을 갖는 DNA 서열을 갖는 유전자에 관한 것이다. 바람직하게는, 상기한 유전자는 서열번호 6의 유전자이다. 또한, 본 발명은 상기한 유전자를 포함하는 재조합 벡터에 관한 것이다. 바

람직하게는, 재조합 벡터에 포함되는 유전자는 서열번호 6의 유전자이다. 바람직하게는, 재조합 벡터는 pYIL-TFP3 또는 pYIL-KRTFP3이다. 또한, 본 발명은 상기한 재조합 벡터로 형질전환된 세포에 관한 것이다. 바람직하게는, 형질전환된 세포는 pYIL-KRTFP3으로 형질전환된 에서리키아 콜라이 DH50/pYIL-KRTFP3(KCTC 10546BP)이다.

<91> 또 다른 양태로서, 본 발명은 서열번호 7로 기재되는 단백질융합인자 TFP-4 단백질, 이의 유사체 또는 이의 단편에 관한 것이다. 또한, 본 발명은 서열번호 7로 기재되는 단백질융합인자 TFP-4 단백질, 이의 유사체 또는 이의 단편을 코딩하는 유전자에 관한 것이다. 달리, 본 발명은 서열번호 7로 기재되는 아미노산 서열 또는 이의 상동성, 바람직하게는 75%, 보다 바람직하게는 85%, 더욱 바람직하게는 90%, 가장 바람직하게는 95% 이상의 상동성을 갖는 아미노산 서열을 갖는 단백질융합인자 TFP-4 단백질에 관한 것이다. 또한, 본 발명은 서열번호 7로 기재되는 단백질융합인자 TFP-4 단백질을 코딩하는 DNA 서열 또는 이의 상동성, 바람직하게는 75%, 보다 바람직하게는 85%, 더욱 바람직하게는 90%, 가장 바람직하게는 95% 이상의 상동성을 갖는 DNA 서열을 갖는 유전자에 관한 것이다. 바람직하게는, 상기한 유전자는 서열번호 8의 유전자이다. 또한, 본 발명은 상기한 유전자를 포함하는 재조합 벡터에 관한 것이다. 바람직하게는, 재조합 벡터에 포함되는 유전자는 서열번호 8의 유전자이다. 바람직하게는, 재조합 벡터는 pYIL-TFP4 또는 pYIL-KRTFP4이다. 또한, 본 발명은 상기한 재조합 벡터로 형질전환된 세포에 관한 것이다. 바람직하게는, 형질전환된 세포는 pYIL-KRTFP4로 형질전환된 에서리키아 콜라이 DH50/pYIL-KRTFP4(KCTC 10547BP)이다.

<92> 본 발명에서 단백질융합인자 단백질 또는 유전자에 대해 사용된 용어 "유사체"란 단백질을 융합인자 유전자를 난발현성 단백질의 유전자와 융합하는 경우 난발현성 단백질의 분비생산을 유도하여 단백질융합인자 활성을 나타내는 작용적 등가물을 의미하며, 단백질융합인자 단백질

의 경우 예를 들어 동일한 성질을 갖는 아미노산끼리의 치환(예: 소수성 아미노산의 다른 소수성 아미노산으로의 치환, 친수성 아미노산의 다른 친수성 아미노산으로의 치환, 염기성 아미노산의 다른 염기성 아미노산으로의 치환, 산성 아미노산의 다른 산성 아미노산으로의 치환), 아미노산의 결실, 삽입 또는 이들의 조합을 포함할 수 있다.

<93> 본 발명에서 단백질을 합인자 단백질 또는 유전자에 대해 사용된 용어 "단편"이란 게놈성 라이브러리 또는 cDNA 라이브러리를 통해 밝혀진 단백질용합인자 유전자의 전체 서열중 일부 서열을 제거하더라도 난발현성 단백질의 분비에 영향을 끼치지 않거나 분비를 촉진시키는 단백질용합인자 유전자 또는 이러한 유전자에 의해 코딩된 단백질을 의미한다.

<94> 본 발명에서 단백질용합인자 단백질 또는 유전자에 대해 사용된 용어 "상동성"이란 야생형(wild type) 아미노산 서열 및 야생형 핵산 서열과의 유사한 정도를 나타내기 위한 것으로서, 단백질의 경우 본 발명의 TFP 단백질의 아미노산 서열과 바람직하게는 75%이상, 보다 바람직하게는 85% 이상, 더욱 바람직하게는 90%이상, 가장 바람직하게는 95% 이상 동일할 수 있는 아미노산 서열을 포함한다. 일반적으로, 단백질 상동물은 목적 단백질과 동일한 활성 부위를 포함할 것이다. 유전자의 경우 본 발명의 TFP 단백질을 코딩하는 DNA 서열과 바람직하게는 75%이상, 보다 바람직하게는 85% 이상, 더욱 바람직하게는 90%이상, 가장 바람직하게는 95% 이상 동일할 수 있는 DNA 서열을 포함한다. 이러한 상동성의 비교는 육안으로나 구입이 용이한 비교 프로그램을 이용하여 수행할 수 있다. 시판되는 컴퓨터 프로그램은 2개 이상의 서열간의 상동성을 백분율(%)로 계산할 수 있으며, 상동성(%)은 인접한 서열에 대해 계산될 수 있다.

<95> 난발현성 단백질의 분비 생산을 위해 본 발명에 따라 밝혀진 단백질용합인자는 난발현성 단백질의 유전자와 융합하여 사용되며 난발현성 단백질의 분비 생산을 위해 벡터에 삽입된

다. 본 발명에서 "벡터"는 적당한 숙주 세포에서 단백질의 발현을 조절할 수 있는 조절 서열 (regulatory sequence)에 작동 가능하도록 연결된 DNA 서열 및 기타 유전자 조작을 용이하게 하거나 단백질의 발현을 최적화하기 위해 도입되는 서열들을 함유하는 DNA 작제물을 의미한다. 그러한 조절 서열에는 전사를 조절하기 위한 프로모터(promoter), 전사를 조절하기 위해 선택 적으로 부가된 오퍼레이터(operator), 적절한 mRNA 리보솜 결합 부위 및 전사/번역의 종료를 조절하는 서열들이 포함된다. 외래 유전자를 삽입하기 위한 벡터로는 플라스미드, 바이러스, 코즈미드 등 다양한 형태의 벡터를 사용할 수 있다. 벡터는 클로닝 벡터 및 발현 벡터를 포함 하며, 클로닝 벡터는 외래 DNA가 삽입되어 복제될 수 있는 플라스미드로서, 형질전환 시 숙주 박테리아 세포로 외래의 DNA를 전달시킨다. 발현 벡터는 통상 외래 DNA의 단편이 삽입된 캐리어로서 일반적으로 이중 가닥의 DNA의 단편을 의미한다. 여기서, 외래 DNA는 숙주 세포에서 천연적으로 발견되지 않는 DNA인 이중 DNA를 의미한다. 발현 벡터는 일단 숙주 세포내에 있으면 숙주 염색체 DNA와 무관하게 복제할 수 있으며 삽입된 외래 DNA가 발현될 수 있다. 당업계 에 주지된 바와 같이, 숙주 세포에서 형질감염 유전자의 발현 수준을 높이기 위해서는 해당 유 전자가 선택된 발현 숙주 내에서 기능을 발휘하는 전사 및 해독 발현 조절 서열에 작동 가능하 도록 연결되어야만 한다.

<96> 단백질융합인자를 포함하는 재조합 벡터로의 형질전환과 관련하여 본원 명세서에 사용된 용어 "형질전환"은 DNA를 숙주로 도입하여 DNA가 염색체의 인자로서 또는 염색체 통합완성에 의해 복제가능하게 되는 것을 의미한다. 본 발명에 따른 형질전환에 사용될 수 있는 숙주 세 포는 원핵 또는 진핵 세포 모두를 포함할 수 있다. DNA의 도입효율이 높고, 도입된 DNA의 발 현효율이 높은 숙주가 통상 사용된다. 세균, 예를 들어 에스케리키아, 슈도모나스, 바실러스, 스트렙토마이세스, 진균, 효모와 같은 주지의 진핵 및 원핵 숙주들, 스포도프테라 프루기페르



다(SF9)와 같은 곤충 세포, CHO, COS 1, COS 7, BSC 1, BSC 40, BMT 10 등의 동물 세포 등이 사용될 수 있는 숙주 세포의 예이다. 바람직하게는 에شري키아 콜라이 (*Escherichia coli*)가 사용될 수 있다.

<97> 본 발명자들은 단백질을 합인자 TFP-1, TFP-2, TFP-3 및 TFP-4 유전자의 단편의 난발현성 단백질의 분비에 끼치는 영향을 조사하였으며, TFP-1의 경우 세린, 알라닌-리치 서열, N-글리코실화 부위 또는 이들 모두가 제거된 유전자를 포함하는 벡터의 경우 난발현성 단백질을 분비하지 못하였다. 반면, 5'-UTR(5'-untranslated region)을 제거하는 경우(pYIL-KRTFP1-4) 난발현성 단백질의 발현율이 3배 이상 증가하였으며, 추가의 3' 말단의 추가의 서열을 제거한 경우(pYIL-KRTFP1-3)에도 난발현성 단백질의 분비를 유도하였다.

<98> 또 다른 양태로서, 본 발명은 서열번호 9로 기재되는 단백질을 합인자 TFP1-3 단백질 또는 이의 유사체에 관한 것이다. 또한, 본 발명은 단백질을 합인자 TFP1-3 단백질 또는 이의 유사체를 코딩하는 유전자에 관한 것이다. 달리, 본 발명은 서열번호 9로 기재되는 아미노산 서열 또는 이의 상동성, 바람직하게는 75%, 보다 바람직하게는 85%, 더욱 바람직하게는 90%, 가장 바람직하게는 95% 이상의 상동성을 갖는 아미노산 서열을 갖는 단백질을 합인자 TFP1-3 단백질에 관한 것이다. 또한, 본 발명은 서열번호 9로 기재되는 단백질을 합인자 TFP1-3 단백질을 코딩하는 DNA 서열 또는 이의 상동성, 바람직하게는 75%, 보다 바람직하게는 85%, 더욱 바람직하게는 90%, 가장 바람직하게는 95% 이상의 상동성을 갖는 DNA 서열을 갖는 유전자에 관한 것이다. 또한, 본 발명은 상기한 유전자를 포함하는 재조합 벡터에 관한 것

이다. 바람직하게는, 재조합 벡터에 포함되는 유전자는 서열번호 9로 기재되는 단백질을 코딩하는 유전자이다. 바람직하게는, 재조합 벡터는 pYIL-KRTFP1-3이다. 또한, 본 발명은 상기한 재조합 벡터로 형질전환된 세포에 관한 것이다. 바람직하게는, 형질전환된 세포는 pYIL-KRTFP1-3으로 형질전환된 에شري키아 콜라이 DH5α/pYIL-KRT1-3(KCTC 10548BP)이다.

<99> 또 다른 양태로서, 본 발명은 서열번호 10으로 기재되는 유전자에 의해 코딩되는 단백질을 코딩하는 유전자 TFP1-4 단백질을 코딩하는 유전자 또는 이의 유사체에 관한 것이다. 또한, 본 발명은 단백질을 코딩하는 유전자 TFP1-4 단백질을 코딩하는 유전자 또는 이의 유사체에 관한 것이다. 달리, 본 발명은 서열번호 10으로 기재되는 유전자에 의해 코딩되는 아미노산 서열 또는 이의 상동성, 바람직하게는 75%, 보다 바람직하게는 85%, 더욱 바람직하게는 90%, 가장 바람직하게는 95% 이상의 상동성을 갖는 아미노산 서열을 갖는 단백질을 코딩하는 유전자 TFP1-4 단백질에 관한 것이다. 또한, 본 발명은 서열번호 10로 기재되는 DNA 서열 또는 이의 상동성, 바람직하게는 75%, 보다 바람직하게는 85%, 더욱 바람직하게는 90%, 가장 바람직하게는 95% 이상의 상동성을 갖는 DNA 서열을 갖는 유전자에 관한 것이다. 또한, 본 발명은 서열번호 10으로 기재되는 단백질을 코딩하는 유전자 TFP1-4을 코딩하는 유전자 또는 이의 유사체를 포함하는 재조합 벡터에 관한 것이다. 달리, 본 발명은 서열번호 10으로 기재되는 단백질을 코딩하는 DNA 서열 또는 이의 상동성, 바람직하게는 75%, 보다 바람직하게는 85%, 더욱 바람직하게는 90%, 가장 바람직하게는 95% 이상의 상동성을 갖는 DNA 서열을 갖는 유전자를 포함하는 재조합 벡터에 관한 것이다. 바람직하게는, 재조합 벡터는 pYIL-KRTFP1-4이다. 또한, 본 발명은 상기한 재조합 벡터로 형질전환된 세포에 관한 것이다. 바람직하게는, 형질전환된 세포는 pYIL-KRTFP1-4으로 형질전환된 에شري키아 콜라이 DH5α/pYIL-KRT1-4(KCTC 10549BP)이다.

<100> 또 다른 양태로서, 본 발명은 서열번호 1, 3, 5, 7 또는 9로 기재되는 단백질, 이의 유사체 또는 이의 단편, 또는 서열번호 10으로 기재되는 유전자, 이의 유사체 또는 이의 단편에 의해 코딩되는 단백질을 이용하여, 난발현성 단백질을 재조합 방법으로 생산하는 방법에 관한 것이다. 달리, 본 발명은 서열번호 1, 3, 5, 7 또는 9로 기재되는 아미노산 서열 또는 이의 상동성, 바람직하게는 75%, 보다 바람직하게는 85%, 더욱 바람직하게는 90%, 가장 바람직하게는 95% 이상의 상동성을 갖는 아미노산 서열을 갖는 단백질, 또는 서열번호 10으로 기재되는 유전자에 의해 코딩되는 아미노산 서열 또는 이의 상동성, 바람직하게는 75%, 보다 바람직하게는 85%, 더욱 바람직하게는 90%, 가장 바람직하게는 95% 이상의 상동성을 갖는 아미노산 서열을 갖는 단백질을 이용하여, 난발현성 단백질을 재조합 방법으로 생산하는 방법에 관한 것이다. 바람직하게는 서열번호 1에 기재되는 단백질은 서열번호 2에 기재되는 유전자에 의해 코딩되며, 서열번호 3에 기재되는 단백질은 서열번호 4에 기재되는 유전자에 의해 코딩되고, 서열번호 5에 기재되는 단백질은 서열번호 6에 기재되는 유전자에 의해 코딩되며, 서열번호 7에 기재되는 단백질은 서열번호 8에 기재되는 유전자에 의해 코딩된다. 바람직하게는, 난발현성 단백질은 인체 인터류킨-2 또는 인체 G-CSF이다.

<101> 이하, 본 발명을 실시예에 의해 상세히 설명한다. 하기 실시예는 본 발명을 예시하는 것일 뿐, 본 발명의 내용이 하기 실시예에 한정되는 것은 아니다.

<102> <실시예 1> 효모 인버테이즈 결여 변이주 제조

- <103> 본 발명자들은 난발현성 단백질에 대한 단백질융합인자를 초고속 선별하기 위해서 효모 인버테이즈를 리포터로 사용하여 수크로즈 배지에서의 세포 성장여부를 이용한 자동선별시스템을 제조하였다.
- <104> 구체적으로 벡터에 포함된 인버테이즈 유전자를 형질전환 후의 선별 리포터유전자로 사용하기 위해서 인버테이즈 활성이 결여된 효모가 필요하며 이를 위해 염색체에 존재하는 *INV2* 유전자를 결실시켰다. 유전자 결실 유도용 카세트 제조를 위해 플라스미드 pRB58 (Carlson 등, Cell, 1982, 20, 145)을 제한효소 *EcoRI* 과 *XhoI*으로 처리하여 *INV2* 코딩유전자를 회수하여 pBluescript KS+의 *EcoRI/XhoI* 부위에 도입하여 pBI Δ BX를 제조하였다. 도 1 에서 보는 바와 같이 pBI Δ BX에 포함된 *INV2* 유전자의 *HindIII-XbaI* site에 190 bp의 반복서열(Tc190)을 양 말단에 포함하는 *URA3* 유전자를 삽입하여 pBIU를 제조하였다. pBIU를 제한 효소 *EcoRI-XhoI*으로 처리한 후 사카로마이세스 세레비시에 Y2805 Δ *gal1* (*Mat a ura3 INV2 pep4::HIS3 gal1 can1*) 균주에 형질전환하여 우라실이 없는 선택배지에서 형질전환체 Y2805 Δ *gal1* Δ *inv2U*(*Mat a inv2::URA3 pep4::HIS3 gal1 can1*)를 선별하였다.
- <105> 선별된 형질전환체의 인버테이즈 활성이 소멸되었는지 여부를 확인하기 위하여 단일 콜로니를 글루코스와 수크로즈를 단일 탄소원으로 공급한 각각의 배지에서 배양한 결과 글루코스에서는 정상적으로 성장하였고 수크로즈에서도 대조구에 비해 성장이 매우 느렸지만 성장할 수 있었다. 그러나 배양 중 배지로 분비되는 인버테이즈의 양을 조사하기 위해서 *INV2+* 균주와 Δ *inv2* 균주를 배양하고 배양 상등액에 존재하는 단백질을 SDS-PAGE로 분리한 후 젤을 수크로즈 용액에서 30분간 반응한 다음 TTC (2,3,5-triphenyl-tetrazolium chloride)로 발색한 자이모그램(zymogram) 분석을 한 결과 (도 2) Δ *inv2* 균주의 경우 대부분의 invertase 활성이 소실되었음을 알 수 있었다. 그러나 수크로즈 배지에서 매우 느리지만 성장하는 문제가 있는데

이는 미토콘드리아의 기능을 통한 글루코네오제네시스(gluconeogenesis)에 의해 세포가 부분적으로 성장하는 것으로 추정되었다. 따라서 이러한 문제를 제거하기 위해 미토콘드리아 전자전달계 작용 억제물질인 안티마이신 A를 첨가한 후 인버테이즈의 발현이 없는 균주의 성장여부를 확인한 결과 균주의 성장을 완전히 억제할 수 있었다(도 3).

<106> 선별된 Y2805 $\Delta gal1\Delta inv2U(Mat a inv2::URA3 pep4::HIS3 gal1 can1)$ 균주에 인체 cDNA 라이브러리를 함유하는 *URA3* 벡터를 재형질전환하기 위해서는 *INV2* 유전자 결실용으로 사용한 *URA3* 유전자를 제거할 필요가 있는데 이를 위해 배양세포를 불화오르티산(5-fluoroorotic acid, 5-FOA)배지에서 배양하여 *URA3* 유전자가 팝-아웃(pop-out)된 균주 Y2805 $\Delta gal1\Delta inv2U(Mat a ura3 inv2::Tc190 pep4::HIS3 gal1 can1)$ 를 선별하였다(도 1). 염색체상의 *INV2* 유전자가 예상한 바와 같이 결실되고 *URA3* 유전자가 다시 팝-아웃되었는지 여부를 서던블롯팅(Southern blotting)을 통해서 확인하였다(도 4). Y2805 균주의 염색체를 *EcoRI*으로 처리하고 *INV2* 유전자를 프로브로 사용할 경우 약 4.3 kb의 절편이 확인되는데 *URA3* 유전자가 삽입된 후(Y2805 $\Delta gal1\Delta inv2U$) 약 5.0 kb로 증가되었다가 *URA3* 유전자가 팝-아웃되면(Y2805 $\Delta gal1\Delta inv2$) 약 3.7 kb로 줄어든다. 결과에서 보는 바와 같이 *INV2* 유전자가 예상한 바와 같이 정확히 결실되었고 *URA3* 유전자가 팝-아웃되었음을 확인하였다.

<107> <실시예 2> 인버테이즈와 융합을 통한 자동선별 시스템의 확인

<108> 인버테이즈 유전자가 결실된 균주(Y2805

$\Delta inv2$)에서 인버테이즈와 융합된 단백질의 발현을 통해 수크로즈 배지에서의 자동선택을 확인하기 위하여 효모에서 발현이 잘 되는 인체단백질인 인체 혈청알부민(human serum albumin, HSA)과 난발현단백질인 인체 인터류킨-2(IL-2)를 이용하였다.

<109> 먼저 본 발명자들은 알부민과 인버테이즈가 융합된 벡터 pGHSA-INV2를 만들기 위하여 *Sfi*I 인식서열을 갖는 중합효소 연쇄반응용 센스 프라이머 JH97(서열번호 11)과 안티센스 프라이머 JH119(서열번호 12)를 사용하였으며 pYHSA5(Kang 등, J. Microbiol. Biotechnol., 1998, 8, 42)를 주형으로 하고 Pfu 중합효소(스트라타진사, 미국)을 이용하여 중합효소 연쇄반응(94℃에서 5분동안 1회; 94℃ 30초간, 55℃ 30초간, 72℃ 2분간 반응을 25회; 72℃에서 7분간 1회)하여 알부민유전자 약 1.8kb의 절편을 얻었다. 또한, 프라이머 JH99(서열번호 13) 및 JH100(서열번호 14)과 주형으로 pRB58을 사용하여 인버테이즈 유전자를 동일한 방법으로 중합효소 연쇄반응하여 회수한 후 제한효소 *Sfi*I/*Sa*I으로 처리하고 *Pst*I/*Sfi*I으로 처리된 알부민 유전자와 함께 *Pst*I/*Sa*I으로 처리된 pBluescript(스트라타진사, 미국)에 삽입하였다. pYHSA5를 제한효소 *Sac*I/*Pst*I으로 처리하여 *GAL* 프로모터 및 알부민유전자 일부를 포함하는 절편을 얻고 앞서 제조된 알부민 유전자 일부 및 인버테이즈 유전자를 함유하는 *Pst*I/*Sa*I 절편을 다시 *Sac*I/*Sa*I 처리된 YEG α-HIR525 벡터(Choi 등, Appl Microbiol Biotechnol., 1994, 42, 587)에 동시에 삽입하여 pGHSA-INV2를 제조하였다. 인터류킨-2와 인버테이즈의 융합발현 벡터 제조를 위해서 인터류킨-2 유전자를 프라이머 JH106(서열번호 15) 및 JH107(서열번호 16)을 사용하여 얻은 유전자를 함유하는 pBKS-IL2를 *Sfi*I으로 처리한 절편과 *GAL* 프로모터 및 *INV* 분비시그날을 함유하는 *Sac*I-*Sfi*I 절편을 *Sac*I/*Sfi*I으로 처리된 pGHSA-INV2에 동시에 삽입하여 pGIL2-INV2를 제조하였다.

<110> 알부민과 인버테이즈의 융합단백질을 발현하는 벡터 pGHSA-INV2와 인터류킨-2와 인버테이즈의 융합단백질을 발현하는 pGIL2-INV2 및 인버테이즈 만을 발현하는 pRB58을 인버테이즈 유전자가 결실되어 수크로즈 배지에서 성장할 수 없는 균주(Y2805 Δ inv2)에 각각 형질전환하고 글루코스를 탄소원으로 첨가한 배지(UD)와 수크로즈를 탄소원으로 첨가한 배지(YPSA)에 도말하여 세포의 성장여부를 관찰하였다(도 5). 인버테이즈를 정상적으로 발현하는 벡터인 pRB58의 경우에는 두 가지 탄소원에서 모두 제대로 성장하였으며, 또한 인버테이즈가 효모에서 발현이 잘되는 단백질인 알부민과 융합된 경우(pGHSA-INV2)에도 역시 두가지 탄소원을 모두 잘 이용하며 성장하였다. 그러나 효모에서 발현되기 어려운 단백질인 인터류킨-2의 경우(pGIL2-INV2)에는 글루코스 배지에서는 정상적으로 성장하지만 수크로즈 배지에서는 전혀 성장하지 못하였다. 이는 예상한 바와 같이 인터류킨-2가 세포내에서 발현이 불가능하여 이에 융합된 인버테이즈도 발현되지 못하기 때문으로 판단되었다. 따라서 인버테이즈 유전자가 결실되어 수크로즈 배지에서 성장할 수 없는 균주(Y2805 Δ inv2)에 형질전환한 인버테이즈 유전자의 발현여부를 이용하여 자동선별이 가능함을 알 수 있었다.

<111> <실시예 3> 난발현단백질 인체 인터류킨-2를 이용한 단백질융합인자 선별 벡터의 제조

<112> 인터류킨-2와 인버테이즈가 융합된 벡터 pGIL2-INV2에 융합단백질의 분비를 유도할 수 있는 적절한 단백질융합인자를 확보하기 위하여 세가지 리딩프레임(reading frame)을 갖는 라이브러리 제조 벡터 pYHTS-F0, F1 및 F2를 제조하였다(도 6). 세 개의 리딩프레임과 *Bam*HI 인식서열을 갖는 중합효소 연쇄반응용 센스 프라이머 JH120(서열번호 17), JH121(서열번호 18)

및 JH122(서열번호 19)와 안티센스 프라이머 JH123(서열번호 20)를 이용하고 pGIL2-INV2를 주형으로 사용하여 Pfu 중합효소(스트라타진사, 미국)를 이용한 중합효소 연쇄반응(94℃에서 3분 동안 1회; 94℃ 30 초간, 55℃ 30초간, 72℃ 1분간 반응을 25회; 72℃에서 7분간 1회)하여 인터류킨-2와 일부의 인버테이즈 유전자를 포함하는 약 1.2 kb의 절편을 각각 얻었다. 다시 JH124(서열번호 21)와 JH95(서열번호 22)를 프라이머로 이용하고 pGIL2-INV2를 주형으로 사용하여 동일한 조건으로 중합효소 연쇄반응하여 일부의 인버테이즈 유전자를 포함하는 약 0.9 kb의 절편을 각각 얻었다. 각각의 유전자를 아가로스젤로부터 회수한 다음 세 가지 리딩프레임을 갖는 1.2 kb 절편 세 종류와 0.9 kb 절편을 각각 혼합하여 센스 프라이머 JH120(서열번호 17), JH121(서열번호 18) 및 JH122(서열번호 19)와 안티센스 프라이머 JH95(서열번호 22)를 이용하여 2차 중합효소 연쇄 반응을 수행하였고 아가로스젤 전기영동을 통해 각각 2.1 kb 절편 세 종류를 회수하였다. 회수된 3종의 2.1kb 절편을 제한효소 *Bam*HI과 *Sa*I로 처리하여 *Bam*HI과 *Sa*I로 처리된 pGIL2-INV2에 각각 결합하여 pYHTS-F0, F1 및 F2를 제조하였다.

<113> <실시예 4> 효모 유전체로부터 맞춤형 단백질융합인자 라이브러리 제조

<114> 단백질융합인자 라이브러리 제조를 위해서 효모 사카로마이세스 세레비시에 및 한세놀라 폴리모르파 유래의 염색체 DNA를 이용하였다. 각각의 염색체를 제한효소

*Sau3AI*으로 부분절단 후 아가로스 젤에서 0.5 내지 1.0 kb 크기의 DNA를 회수한 다음 *BanHI* 으로 절단하고 송아지 장 유래의 포스파테이즈(calf intestine phosphatase)로 처리한 pYHTS-F0, F1 및 F2 혼합 벡터에 연결하였다(도 6). 연결된 DNA를 에شري키아 콜라이 DH5 α 에 형질전환하고 엠피실린이 포함된 LB배지(1% Bacto-tryptone, 0.5 % 효모 추출물, 1% NaCl)에 도말한 후 37°C에서 하루 배양하였다. 각각의 염색체로부터 제조된 라이브러리 DNA를 이용하여 약 5×10^4 세포의 형질전환체 라이브러리를 확보하였다. 모든 형질전환체를 멸균증류수를 이용하여 회수하고 회수된 균체로부터 라이브러리 DNA를 플라스미드 추출 키트 (바이오니어)를 이용하여 분리하였다.

<115> <실시예 5> 난분비성 단백질 인체 인터류킨-2에 적합한 맞춤형 단백질융합인자 자동선별

<116> 상기한 방법으로 제조된 라이브러리 DNA를 효모 사카로마이세스 세레비시에 Y2805 $\Delta gal1 \Delta inv2$ (*Mat a ura3 inv2::Tc190 pep4::HIS3 gal1 can1*)에 리튬 아세테이트 방법(Hills 등, Nucleic acids Res. 1991, 19, 5791)으로 형질전환한 다음 우라실이 결핍된 UD 최소배지(0.67% 아미노산이 결핍된 효모 기질, 적정농도의 각종 아미노산 혼합체, 2% 글루코스)와 수크로즈 및 안티마이신 A를 포함하는 복합배지 YPGSA(1% 효모추출물, 2% 펩톤, 2% 수크로즈, 0.03% 갈락토즈, 1 μ g/mL 안티마이신 A)에 각각 도말하여 30°C에서 5일간 배양하였다. 배양 후 각 배지에서 형성된 균체의 수는 표 1(단백질융합인자 도입 전후의 형질전환체 수의 비교)에 표시한 바와 같다. 라이브러리 제작을 위해 사용한 벡터 (pYHTS-F0, F1 및 F2)만을 형질전환한 경우에는 글루코스 배지에서는 약 1×10^4 정도의 균체가 형성되었으나 탄소원으로 수크로즈를 사용

한 배지에서는 예상한 바와 같이 전혀 성장하지 못하였다. 그러나 효모 유전체 라이브러리를 삽입한 경우에는 약 11개의 형질전환체가 성장하여 삽입된 단백질을 합인자의 도움으로 인버테이즈가 분비되었음을 예상할 수 있었다.

<117> 【표 1】

형질전환 균주	형질전환체수	
	UD 최소배지 (글루코스)	YPSGA (수크로스)
Y2805 $\Delta gal1 \Delta inv2$ (숙주세포)	0	0
pYHTS vector	$\sim 1 \times 10^4$	0
pYHTS+genomic library(<i>S. cerevisiae</i>)	$\sim 1 \times 10^4$	10
pYHTS+genomic library(<i>H. polymorpha</i>)	$\sim 1 \times 10^4$	1

<118> <실시예 6> 단백질을 합인자의 분석

<119> 수크로스 배지에서 성장한 각 균체를 YPD(1% 효모추출물, 2% 펩톤, 2% 글루코스)배지에서 24시간 배양한 후 회수된 세포를 파쇄한 다음 도입된 플라스미드를 분리하고 에서리키아 콜라이에 재형질전환하였다. 형질전환된 에서리키아 콜라이으로부터 플라스미드를 분리한 후 제한효소 처리하여 삽입된 유전자의 유무를 확인하였고 염기서열 분석을 통해 4 종류의 서로 다른 서열을 갖는 유전자가 삽입되어 단백질을 합인자 역할을 하는 것을 알 수 있었다(표 2: 수크로스배지에서 성장한 균체로부터 분리된 플라스미드에 삽입된 신규 단백질을 합인자).

<120> 【표 2】

플라스미드	융합인자	효모유전자	융합된 아미노산의 수 (전체 아미노산의 수)	특징
pYHTS-TFP1	<i>TFP-1</i>	<i>Yar066w</i>	105(203)	PRE, N-gly, Ser-rich, GPI
pYHTS-TFP2	<i>TFP-2</i>	<i>Yar026c</i>	117(169)	PRE, N-gly
pYHTS-TFP3	<i>TFP-3</i>	<i>Yj1158c</i>	104(227)	PRE-PRO, O-gly, PIR
pYHTS-TFP4	<i>TFP-4</i>	<i>Unknown</i>	50(unknown)	PRE

<121> 1. 단백질융합인자 1 (translational fusion partner 1, TFP-1)

<122> 본 발명자들은 인터류킨-2와 인버테이즈 융합단백질을 효율적으로 세포 밖으로 분비시킬 수 있는 단백질융합인자 1(TFP-1)(서열번호 2)을 발굴하였다. TFP-1은 효모 사카로마이세스 세레비지에 유전자 *Yar066w*와 동일한 유전자이며 α 1,4-글루칸 글루코시다제 (STA1)와 유사하며 글리코실포스파티딜이노시톨(GPI)-앵커 (glycosyl- phosphatidylinositol(GPI)-anchor)를 함유한 단백질로 아직 기능이 알려지지 않은 유전자이다. 인터류킨과 융합된 아미노산의 수는 전체 203개 중 105개이며 단백질 분비를 위한 시그널로 23개 아미노산의 분비 시그널을 함유하고 있으며 N-글리코실화 부위와 세린, 알라닌-리치(serine-rich) 서열을 함유하고 있다.

<123> 2. 단백질융합인자 2 (TFP-2)

<124> 본 발명자들은 인터류킨-2와 인버테이즈 융합단백질을 효율적으로 세포 밖으로 분비시킬 수 있는 단백질융합인자 2(TFP-2)(서열번호 4)를 발굴하였다. TFP-2은 효모 사카로마이세스 세레비지에 유전자 *Yar026c*와 동일한 유전자이며 아직 기능이 알려지지 않은 유전자이다. 인터류킨-2와 융합된 아미노산의 수는 전체 169개 중 117개이며 단백질 분비를 위한 시그널로 19개 아미노산의 분비 시그널을 함유하고 있으며 3개의 N-글리코실화 부위를 함유하고 있다.

<125> 3. 단백질융합인자 3 (TFP-3)

<126> 본 발명자들은 인터류킨-2와 인버테이즈 융합단백질을 효율적으로 세포 밖으로 분비시킬 수 있는 단백질융합인자 3(TFP-3)(서열번호 6)을 발굴하였다. TFP-3은 효모 사카로마이세스 세레비지에 유전자 Yj1158c(CIS3)와 동일한 유전자이며 세포벽에 공유결합된 O-만노실화된 단백질이며 세포분열에 관여하는 *Cik1* 결여변이주에 대한 멀티카피 서프레서(multicopy suppressor)로 알려진 유전자이다. 인터류킨-2와 융합된 아미노산의 수는 전체 227개 중 104개이며 단백질 분비를 위한 시그널로 23개 아미노산의 프리(pre) 분비 시그널과 41개의 프로(pro) 분비서열을 함유하고 있고 Lys-Arg로 구성된 Kex2p 절단서열을 함유하고 있으며 PIR 반복서열을 가지고 있다.

<127> 4. 단백질융합인자 4 (TFP-4)

<128> 본 발명자들은 인터류킨-2와 인버테이즈 융합단백질을 효율적으로 세포 밖으로 분비시킬 수 있는 단백질융합인자 4(TFP-4)(서열번호 8)를 발굴하였다. TFP-4은 효모 한세놀라 폴리모르파 유래의 유전자이며 기능은 알려진 바 없으나 인터류킨-2와 융합된 아미노산의 수는 50개이며 이 중 18개의 아미노산으로 구성된 단백질 분비시그널을 포함하고 있다.

<129> <실시예 7> 배지로 분비된 융합단백질의 분석

<130>

수크로즈 배지에서 성장한 균체가 분비하는 단백질을 분석하기 위해서 상기한 4가지 단백질융합인자를 함유하는 균체를 YPDG 배지(1% 효모추출물, 2% 펩톤, 2% 글루코스, 0.03% 갈락토즈)에서 40시간 배양한 후 세포를 제거하고 남은 배양 상등액에 녹아있는 전체 단백질을 아세톤(최종농도 40%)으로 침전시킨 후 SDS-PAGE 하였으나 인버테이즈가 융합된 상태에서는 인버테이즈 부분의 과도한 당쇄부가로 인해서 단일 단백질 밴드를 확인하기 힘든 문제가 있었기 때문에 각 벡터(pYHTS-TFP1, 2, 3 및 4)에서 인버테이즈를 제거하고 인터류킨-2 유전자에 번역종결코돈을 삽입하였다. 이를 위해 프라이머 JH132(서열번호 23) 및 JH137(서열번호 24)을 이용하여 각각 중합효소 연쇄반응한 후 회수된 유전자 절편을 *SacI*/*SaI*로 처리하고 *SacI*/*SaI*로 처리된 YEG α-HIR525 벡터에 삽입하여 pYIL-TFP1, 2, 3 및 4를 제조하였다. 제조된 4종의 IL-2 발현 벡터를 효모에 형질전환한 후 단일 콜로니를 확보하고 상기한 바와 같은 방법으로 배양하여 배양상등액에 분비된 단백질을 SDS-PAGE하여 분석하였다(도 7). 결과에서 보는 바와 같이 pYIL-TFP2를 제외한 각각의 벡터를 함유한 균체의 배양 상등액에서 단백질의 크기가 서로 다른 강한 밴드들이 관찰되었으며 이들은 IL-2 항체를 이용한 웨스턴브롯팅(도 7)에서 밴드로 확인되어 각각의 분비유도 융합단백질과 IL-2가 융합된 상태임을 확인할 수 있었다. 그러나 각각의 단백질융합인자 및 IL-2 유전자 크기로부터 유추되는 단백질 크기보다 SDS-PAGE 상에서 나타난 실제 크기가 상당한 차이가 있었는데 이는 당쇄부가로 인한 차이일 것으로 추정되어 *N*-글리코실화에 의해서 단백질에 부가된 당을 제거하기 위해서 각 단백질에 엔도에이치(Endo-H) 효소를 처리한 후 다시 SDS-PAGE 분석하였다(도 8). 단백질 서열상 TFP-1에는 한 개의 *N*-글리코실화 유발 서열(28번 내지 30번째 아미노산)이 있으며 TFP-3은 *O*-글리코실화 유도서열을 함유하고 있으며 TFP-4에는 당쇄부가 유도서열이 없었다. 예상한 바와 같이 pYIL-TFP1의 경우에는 엔도에이치 처리 후 분자량이 상당히 감소함을 알 수 있었다. 따라서 pYIL-TFP1의 경우에

는 발현된 단백질이 *N*-글리코실화 되어 있음을 확인할 수 있었으나 pYIL-TFP3의 경우에는 *O*-글리코실화에 의한 당쇄부가이기 때문에 엔도에이치 처리하여도 분자량의 변화가 없었고 pYIL-TFP4의 경우에도 아무런 변화가 없었다.

<131> <실시예 8> 세포내에서 Kex2p 절단에 의한 원형(authentic) 단백질 생산

<132> 본 발명자들은 실시예 7에서 사용한 벡터가 각 단백질 융합인자와 IL-2가 융합된 상태로 배지로 분비되기 때문에 이를 인체 인터류킨-2와 동일한 상태의 원형(authentic)단백질을 생산하기 위해서 세포내에서 단백질융합인자를 자동으로 제거 할 수 있도록 융합인자와 IL-2 사이에 효모가 자체적으로 생산하는 단백분해효소 Kex2p가 인식할 수 있는 절단 부위 (Leu-Asp-Lys-Arg)를 각각 삽입하였다. pYIL-TFP1에 Kex2p 절단서열을 넣기 위해서 pYIL-TFP1을 주형으로 하고 프라이머 JH132(서열번호 23) 및 HY22(서열번호 25), HY23(서열번호 26) 및 JH137(서열번호

24)을 이용하여 각각 중합효소 연쇄반응하고 회수된 2개의 유전자절편을 주형으로 하고 JH132(서열번호 23) 및 JH137(서열번호 24)을 이용하여 2차 중합효소 연쇄반응한 후 얻어진 절편을 제한효소 *SacI*/*SaI*으로 처리하고 *SacI*/*SaI*으로 처리된 YEG α -HIR525 벡터에 삽입하여 pYIL-KRTFP1을 제조하였다. 또한 pYIL-TFP2에 Kex2p 절단서열을 넣기 위해서 pYIL-TFP2를 주형으로 하고 프라이머 JH132(서열번호 23) 및 HY20(서열번호 27), HY21(서열번호 28) 및 JH137(서열번호 24)을 이용하여 각각 중합효소 연쇄반응하고 회수된 2개의 유전자절편을 주형으로 하고 JH132(서열번호 23) 및 JH137(서열번호 24)을 이용하여 2차 중합효소 연쇄반응한 후 얻어진 절편을 제한효소 *SacI*/*SaI*으로 처리하고 *SacI*/*SaI*으로 처리된 YEG α -HIR525 벡터에 삽입하여 pYIL-KRTFP2를 제조하였다. 또한 pYIL-TFP3에 Kex2p 절단서열을 넣기 위해서 pYIL-TFP3을 주형으로 하고 프라이머 JH132(서열번호 23) 및 HY17(서열번호 38), HY18(서열번호 39) 및 JH137(서열번호 24)을 이용하여 각각 중합효소 연쇄반응하고 회수된 2개의 유전자절편을 주형으로 하고 JH132(서열번호 23) 및 JH137(서열번호 24)을 이용하여 2차 중합효소 연쇄반응한 후 얻어진 절편을 제한효소 *SacI*/*SaI*으로 처리하고 *SacI*/*SaI*으로 처리된 YEG α -HIR525 벡터에 삽입하여 pYIL-KRTFP3를 제조하였다. 또한 pYIL-TFP4에 Kex2p 절단서열을 넣기 위해서 pYIL-TFP4를 주형으로 하고 프라이머 JH132(서열번호 23) 및 HY24(서열번호 29), HY25(서열번호 30) 및 JH137(서열번호 24)을 이용하여 각각 중합효소 연쇄반응하고 회수된 2개의 유전자절편을 주형으로 하고 JH132(서열번호 23) 및 JH137(서열번호 24)을 이용하여 2차 중합효소 연쇄반응한 후 얻어진 절편을 제한효소 *SacI*/*SaI*으로 처리하고 *SacI*/*SaI*으로 처리된 YEG α -HIR525 벡터에 삽입하여 pYIL-KRTFP4를 제조하였다.

<133> 제조된 4가지 벡터 중 pYIL-KRTFP1, pYIL-KRTFP3 및 pYIL-KRTFP4를 각각 효모 2805 Δ *gal1 Δ inv2* 균주에 도입하여 단일콜로니를 선별하여 YPDG 배지(1% 효모추출물, 2% 펩톤, 2% 글

루코스, 0.03% 갈락토즈)에서 40시간 배양한 후 세포를 제거하고 남은 배양 상등액을 SDS-PAGE 한 결과 도 9 에서 보는 바와 같이 인체 인터류킨-2와 동일한 크기의 단백질이 분비생산됨을 확인하였다. 인체 인터류킨-2를 분비 생산하는 3가지 단백질융합인자 중 TFP-1이 가장 효율적으로 원형의 인터류킨-2를 분비 생산할 수 있음을 확인하였다.

<134> 상기 벡터 pYIL-KRTFP1, pYIL-KRTFP2, pYIL-KRTFP3 및 pYIL-KRTFP4는 2003년 11월 11일자로 국제기탁기관인 대전시 유성구 어은동 52번지 소재의 KCTC(Korean Collection for Type Cultures)에 각각 수탁번호 KCTC 10544BP, 10545BP, 10546BP 및 10547BP로 기탁하였다.

<135> <실시예 9> 단백질융합인자 1 (TFP-1)의 특성분석

<136> 상기한 바와 같이 IL-2를 가장 효율적으로 분비 생산하는 단백질융합인자로 TFP-1을 선별하여 TFP-1 서열에 존재하는 분비시그널 (a), N-글리코실화 부위 (b), 및 세린, 알라닌-리치 서열 (c), 추가서열 (d) 및 5'-UTR(5'-비번역 서열) 서열(e) 중 어떤 서열이 난분비단백질인 인터류킨-2의 분비에 절대적인 영향을 주는지 확인하기 위하여 도 10에서 보는 바와 같이 TFP-1의 각 특징서열을 하나씩 제거한 결여 유전자를 제조하였다. 먼저 TFP-1에서 서열상 별 특징이 없는 추가서열 d를 제거하기 위해 pYIL-KRTFP1을 주형으로 하고 프라이머 JH143(서열번호 31) 및 JH132(서열번호 23)를 사용하여 중합효소 연쇄반응하였다. 획득된 DNA 절편은 GAL 프로모터와 TFP-1에서 서열 d 가 제거된 절편인 TFP1-1을 포함하고 있다. TFP-1에서 추가서열 d 및 세린, 알라닌-리치 서열(c)을 제거하기 위해 pYIL-KRTFP1을 주형으로 하고 프라이머 JH142(서열번호 32) 및 JH132(서열번호 23)를 사용하여 중합효소 연쇄반응하였으며 획득된 절

편은 *GAL* 프로모터와 서열 c 및 d 가 제거된 절편 TFP1-2를 포함하고 있다. 또한 TFP-1에서 d, c 및 N-글리코실화 부위 (b)를 제거하기 위해 pYIL-KRTFP1을 주형으로 하고 프라이머 JH141(서열번호 33) 및 JH132(서열번호 23)를 사용하여 중합효소 연쇄반응하였으며 획득된 절편에는 *GAL* 프로모터 및 서열 c, d 및 b가 제거된 절편을 TFP1-3를 포함하고 있다. Kex2p 절단부위를 포함하는 IL-2 유전자를 제조하기 위해서 pYIL-KRTFP1을 주형으로 하고 프라이머 JH140(서열번호 34) 및 JH137(서열번호 24)을 이용하여 중합효소 연쇄반응하였다. 회수된 IL-2 절편을 제한효소 *SpeI* 및 *SaI*로 처리한 절편과 앞서 획득된 3가지 절편(TFP1-1, 2, 및 3)을 각각 제한효소 *SacI* 및 *XbaI*으로 처리하고 *SacI* 및 *SaI*으로 처리된 YEG α -HIR525에 각각 삽입하여 도 10에서 보는 바와 같이 pYIL-KRT1-1, pYIL-KRT1-2, 및 pYIL-KRT1-3를 제조하였다. TFP-1의 5'-UTR을 제거하기 위해서 pYIL-KRTFP1을 주형으로 하고 프라이머 HY38(서열번호 35) 및 JH137(서열번호 24)을 이용하여 중합효소 연쇄반응하였으며 회수된 유전자를 *BanHI*/*SaI*으로 처리한 후 *SacI*/*BanHI*으로 처리된 *GAL10* 프로모터와 같이 *SacI*/*SaI*으로 처리된 YEG α -HIR525에 삽입하여 pYIL-KRT1-4를 제조하였다(도 10).

<137> 제조된 4종의 플라스미드 pYIL-KRT1-1, pYIL-KRT1-2, pYIL-KRT1-3 및 pYIL-KRT1-4를 효모에 형질전환하고 단일콜로니를 배양하여 배양 상등액을 SDS-PAGE 분석한 결과 도 11에서 보는 바와 같이 분비시그널, N-글리코실화 및 세린, 알라닌-리치 서열을 모두 포함하고 있는 pYIL-KRT1-3의 경우에만 IL-2 밴드가 확인되었고 세린, 알라닌-리치 서열이 제거된 pYIL-KRT1-2 및 세린, 알라닌-리치 서열 및 N-글리코실화 서열이 제거된 pYIL-TFP1-1의 경우에는 전혀 밴드가 확인되지 않았다. 따라서 IL-2 분비를 효과적으로 유도하기 위해서는 TFP-1에 존재하는 세 가지 특징적인 서열(분비시그널, N-글리코실화 및 세린, 알라닌-리치 서열)이 모

두 필요함을 알 수 있었다. 또한 TFP1의 5'-UTR을 제거한 경우 발현율이 약 3배 이상 증가함을 확인하였다.

<138> 상기 벡터 pYIL-KRTFP1-3 및 pYIL-KRTFP1-4는 2003년 11월 11일자로 국제기탁기관인 대전시 유성구 어은동 52번지 소재의 KCTC(Korean Collection for Type Cultures)에 각각 수탁번호 KCTC 10548BP 및 10549BP로 기탁하였다.

<139> <실시예 10> 단백질융합인자 TFP-1을 이용한 인체콜로니 자극인자(G-CSF)의 분비생산

<140> 난분비 단백질인 인체 인터류킨-2를 효율적으로 분비 생산할 수 있는 단백질융합인자 TFP-1이 다른 인체 난분비 단백질의 분비에도 효과가 있는지 여부를 확인하기 위하여 난분비단백질인 인체 콜로니 자극인자(G-CSF)를 TFP-1에 융합하여 효모에서 발현 및 분비를 확인하였다. 인체 콜로니 자극인자 유전자는 인체 cDNA 라이브러리로부터 프라이머 JH144(서열번호 36) 및 JH145(서열번호 37)를 이용하여 중합효소 연쇄반응을 통해 확보하였으며 확보된 유전자를 제한효소 *Xba*I/*Sa*I로 처리하고 pYIL-KRTFP1의 *Xba*I/*Sa*I 부위에 삽입하여 pYGCSF-TFP1을 제조하였다. 대조구로서 외래단백질 분비생산을 위해서 효모에서 가장 흔히 사용하고 있는 분비 시그널인 교배인자 알파(mating factor alpha)유래의 분비시그널인 MF α 를 사용하였다. *Xba*I/*Sa*I로 처리된 G-CSF 유전자를 YEG α -HIR525의 *Xba*I/*Sa*I 부위에 삽입하여 pYGCSF-MF α 를 제조하였다.

<141> 인체 콜로니자극인자를 발현하기 위해 제조된 2가지 벡터 pYGCSF-TFP1 및 pYGCSF-MF α 를 효모에 형질전환하고 단일콜로니를 분리하여 배양한 후 배양상등액을 SDS-PAGE 하고 G-CSF 항

체를 이용한 웨스턴블롯팅 결과는 도 12 에서 보는 바와 같다. 대조구로 사용한 MF α 분비시 그날의 경우 전혀 G-CSF 밴드가 확인되지 않았으나 TFP-1의 경우에는 G-CSF 밴드가 확인되었다. G-CSF에 대한 항체(케미콘, 미국)를 이용하여 웨스턴블롯팅한 결과 예상한 밴드가 확인됨으로써 단백질융합인자 TFP-1에 의해 난분비성 단백질인 G-CSF도 효과적으로 분비됨을 알 수 있었다. 따라서 TFP-1은 IL-2 또는 G-CSF 이외의 다양한 난분비성 단백질의 분비를 촉진할 수 있는 단백질융합인자로 매우 유용할 것으로 판단되었다.

【발명의 효과】

- <142> 본 발명에 따라 재조합 생산이 어려운 난발현성 단백질을 발현 및 분비 생산할 수 있는 맞춤형 단백질융합인자를 다양한 유전자원으로부터 초고속으로 선별할 수 있다.
- <143> 이러한 기술은 현재까지 재조합 생산이 불가능하거나 발현율이 낮았던 다양한 단백질을 맞춤형 TFP 초고속 발굴 및 이용을 통하여 경제적으로 대량생산이 가능하도록 하는 새로운 기술을 제공한다.

【특허청구범위】

【청구항 1】

- (1) 난발현성 단백질의 유전자(X)와 인프레임(in-frame)으로 연결된 자동선별용 리포터 유전자(R)의 융합 유전자(X-R)를 포함하는 자동 선별 벡터를 제조하는 단계;
- (2) 난발현성 융합 단백질(X-R)의 분비를 유도하는 단백질융합인자(TFP: translational fusion partner)를 포함한 계놈성 DNA 또는 cDNA를 자동 선별 벡터와 연결하여 단백질융합인자 라이브러리를 제조하는 단계;
- (3) 단백질융합인자 라이브러리로 리포터 유전자의 활성이 없는 세포를 형질전환시켜 리포터 단백질의 활성을 검출하는 단계; 및
- (4) 리포터 단백질의 활성을 나타내는 형질전환 세포로부터 유전자를 분리하여 단백질융합인자의 특성을 분석하는 단계를 포함하여, 단백질융합인자를 선별하는 방법.

【청구항 2】

제1항에 있어서, 리포터 유전자가 인버테이즈, 아밀레이즈, 글루코아밀레이즈, 갈락토시데이즈, 수크레이즈, 셀룰레이즈, 자일라네이즈 및 말테이즈중에서 선택되는 유전자인 방법.

【청구항 3】

제1항에 있어서, 난발현성 단백질이 사이토카인, 혈청단백질, 면역글로불린, α -, β - 및 γ -인터페론, 콜로니자극인자, 혈소판 유도된 성장 인자(PDGF), 포스포리파제-활성화 단백질(PLAP), 인슐린, 종양 괴사 인자(TNF), 성장 인자, 호르몬, 칼시토닌(calcitonin), 칼시토닌 유전자 관련 펩타이드, 엔케팔린(enkephalin), 소마토메딘, 에리스로포이에틴, 시상하부 분비 인자, 프롤락틴, 만성 고나도트로핀, 조직 플라스미노겐 활성화제, 성장호르몬 분비 펩타이드

(GHPR), 흥선 체액성 인자(THF), 탄수화물-특이적 효소, 단백질분해 효소, 산화환원 효소, 트랜스퍼라제, 하이드롤라제, 라이아제, 이소머라제 및 리가제중에서 선택되는 방법.

【청구항 4】

제1-2항에 있어서, 난발현성 단백질이 인체 인터류킨-2(IL-2) 또는 인체 콜로니자극인자(G-CSF)인 방법.

【청구항 5】

제1항에 있어서, 게놈성 DNA 또는 cDNA가 캔디다(*Candida*), 디베리오마이세스(*Debaryomyces*), 한세놀라(*Hansenula*), 클루이베로마이세스(*Kluyveromyces*), 피키아(*Pichia*), 스킨조사카로마이세스(*Schizosaccharomyces*), 야로이아(*Yarrowia*), 사카로마이세스(*Saccharomyces*), 아스퍼질러스(*Aspergillus*), 페니실리엄(*Penicillium*), 라이조퍼스(*Rhizopus*) 및 트리코더마(*Trichoderma*)중에서 유래되는 방법.

【청구항 6】

제1항에 있어서, 리포터 유전자의 활성이 없는 세포가 캔디다(*Candida*), 디베리오마이세스(*Debaryomyces*), 한세놀라(*Hansenula*), 클루이베로마이세스(*Kluyveromyces*), 피키아(*Pichia*), 스킨조사카로마이세스(*Schizosaccharomyces*), 야로이아(*Yarrowia*), 사카로마이세스(*Saccharomyces*), 아스퍼질러스(*Aspergillus*), 페니실리엄(*Penicillium*), 라이조퍼스(*Rhizopus*) 및 트리코더마(*Trichoderma*)중에서 선택되는 방법.

【청구항 7】

제1항에 있어서, 선별 벡터가 프로모터 유전자, 번역개시 및 종결코돈이 제거된 난발현성 단백질을 코딩하는 유전자 및 이에 인프레임으로 융합된 인버테이즈 코딩 유전자를 포함하는 방법.

【청구항 8】

제1항에 있어서, 선별 벡터에 포함되는 프로모터가 *GAPDH*, *PGK*, *ADH*, *PHO5*, *GAL1* 및 *GAL10*중에서 선택되는 방법.

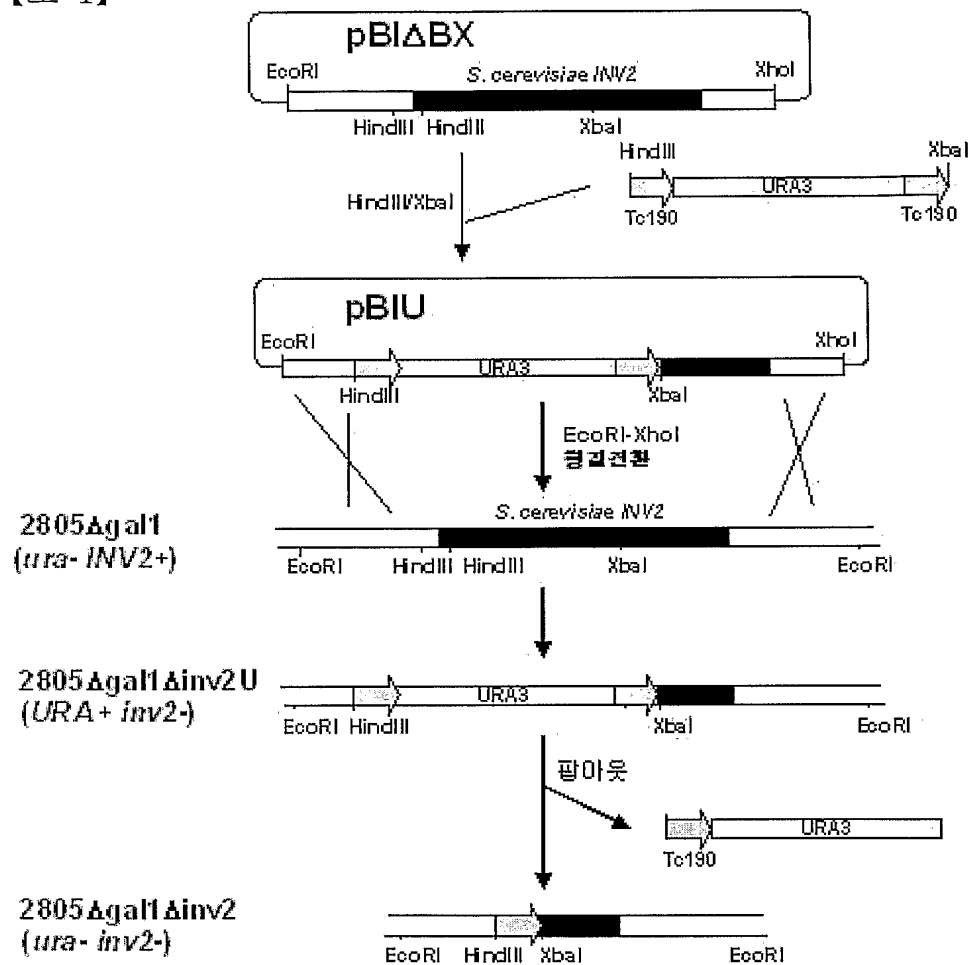
【청구항 9】

- (1) 인버테이즈 유전자인 INV2(I)를 결실시킨 효모변이주를 제조하는 단계;
 - (2) 효모 *GAL10* 프로모터에 의해서 발현이 조절되며 인버테이즈(I) 유전자와 난발현성 유전자(X)가 인프레임(in-frame)으로 융합된 유전자(X-I)를 함유하는 효모 HTS (high throughput selection) 벡터 pYHTS를 제조하는 단계;
 - (3) 인버테이즈와 난발현성 단백질의 융합 유전자(X-I)를 분비시킬 수 있는 효모 유전체로부터 단백질융합인자(TFP: translational fusion partner) 라이브러리를 pYHTS 벡터에 제조하는 단계;
 - (4) 제조된 라이브러리를 단계 (1)에서 제조된 효모에 형질전환하고 수크로스를 단일탄소원으로 포함하고 있는 배지에서 자동선별하는 단계;
 - (5) 수크로스 배지에서 성장한 효모를 배양하여 배지로 분비된 단백질을 확인하는 단계;
- 및

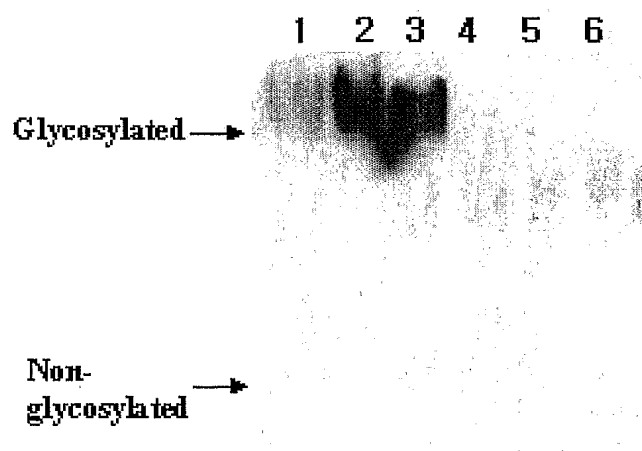
(6) 효모로부터 유전자를 분리하여 단백질융합인자 특성을 분석하는 단계를 포함하여,
단백질융합인자를 선별하는 방법.

【도면】

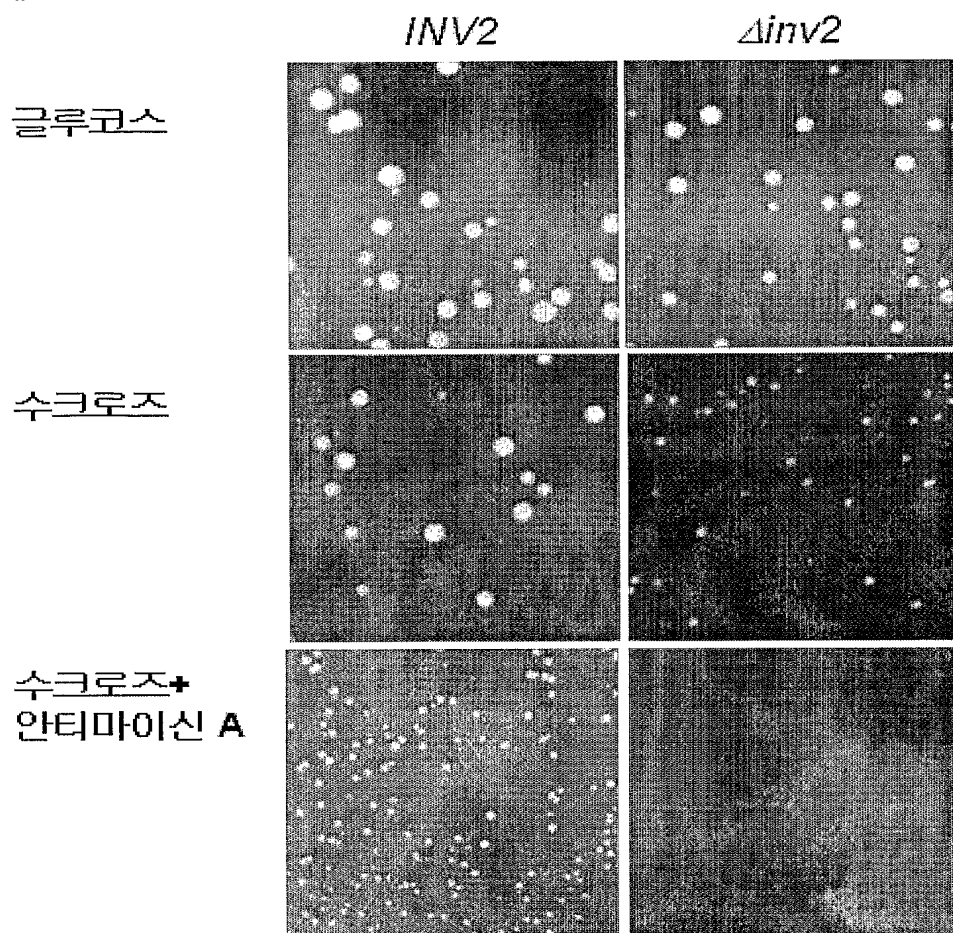
【도 1】



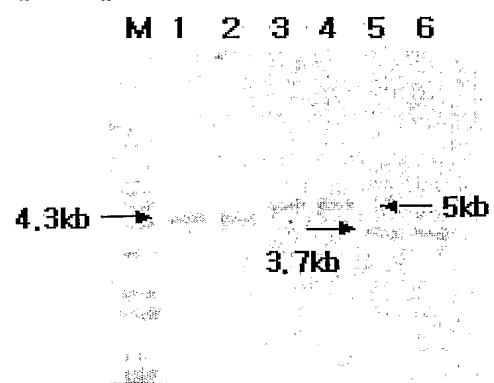
【도 2】



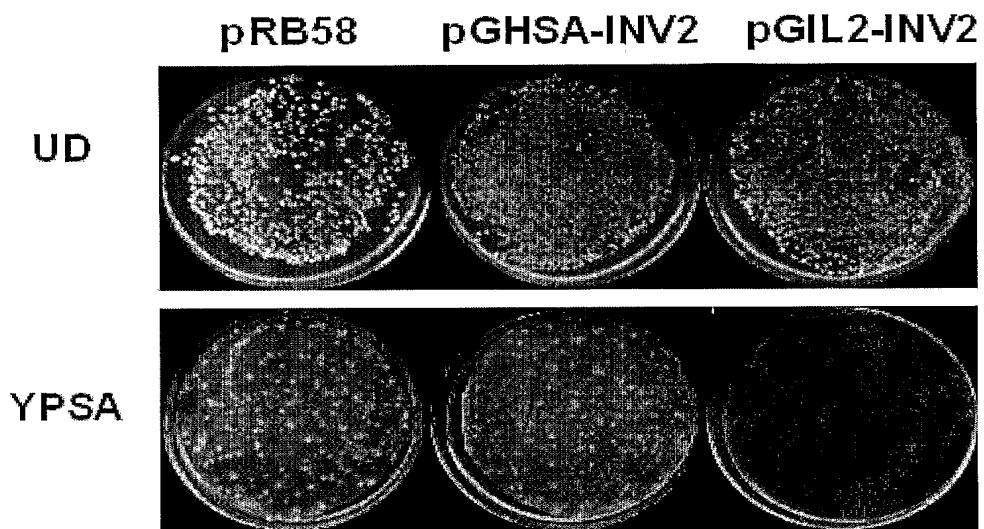
【도 3】



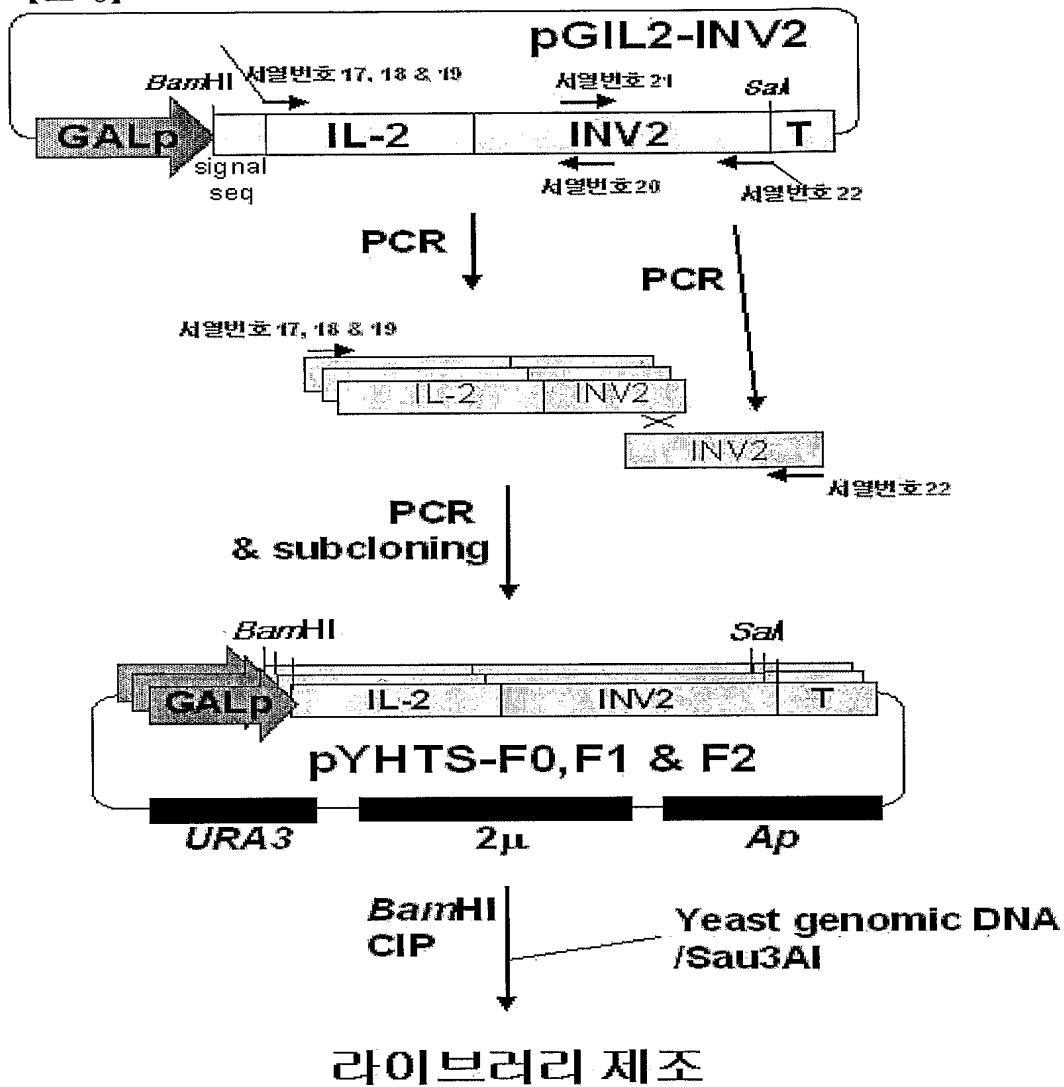
【도 4】



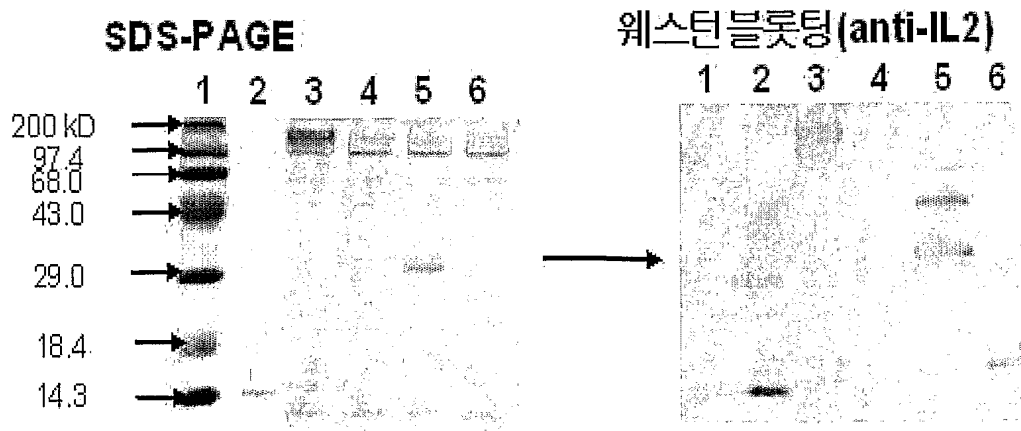
【도 5】



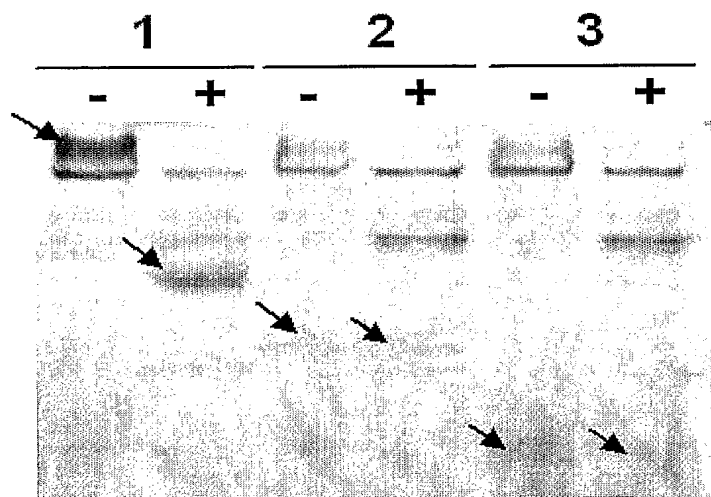
【도 6】



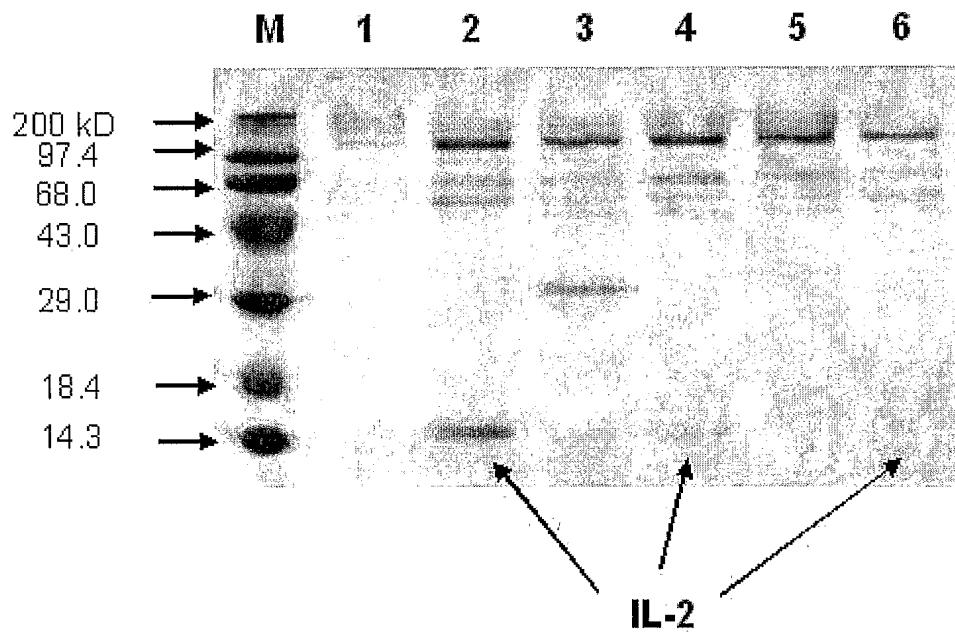
【도 7】



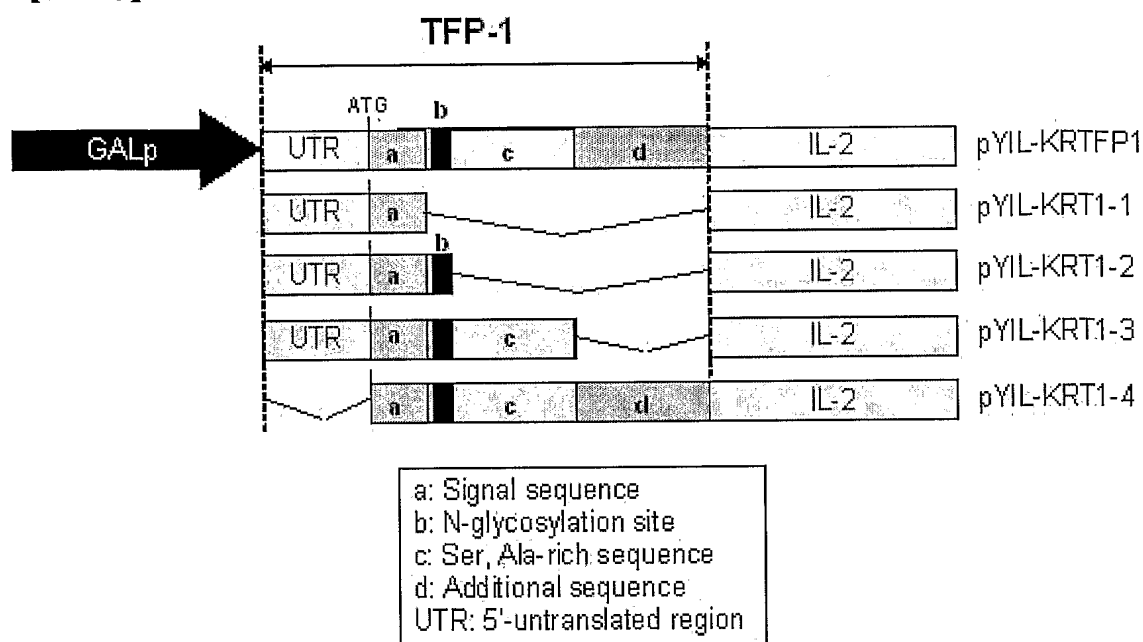
【도 8】



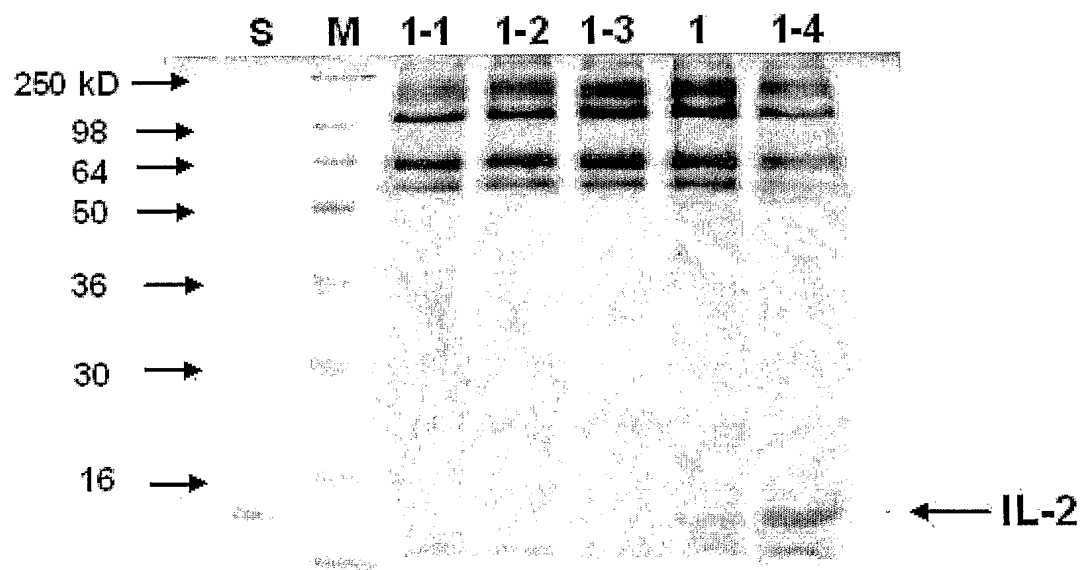
【도 9】



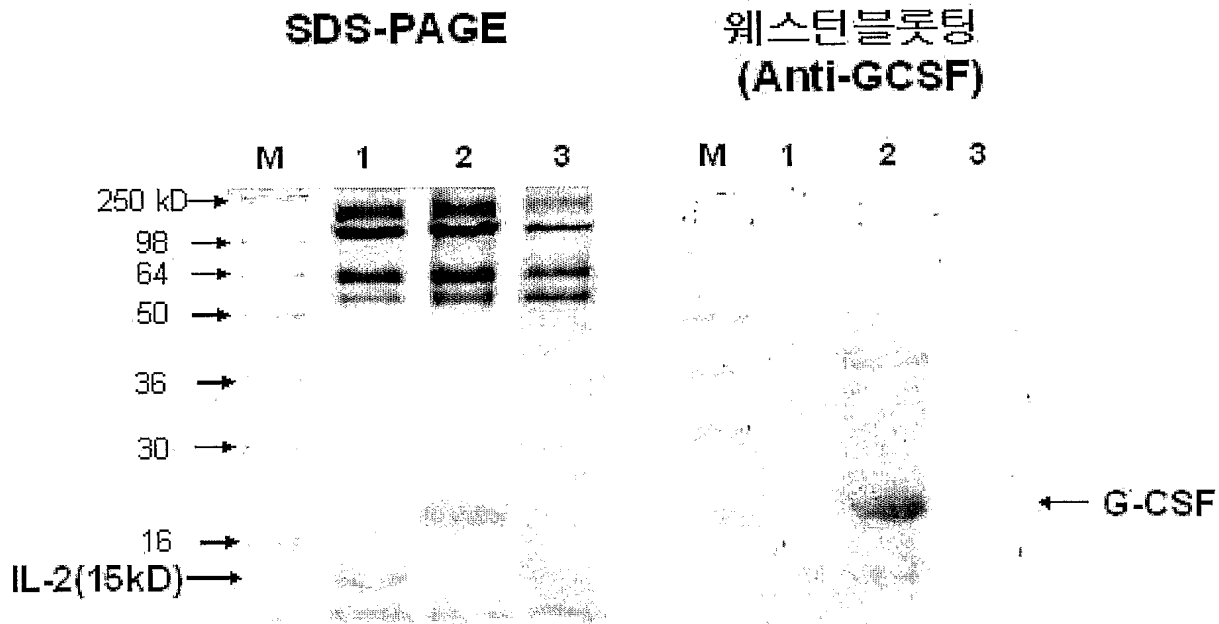
【도 10】



【도 11】



【도 12】



【서열목록】

<110> Korea Research Institute of Bioscience and Biotechnology <120> Rapid
screening method of suitable translational fusion partners for producing
recombinant proteins <160> 39 <170> KopatentIn 1.71 <210> 1 <211> 105 <
212> PRT <213> Saccharomyces cerevisiae <400> 1 Met Phe Asn Arg Phe Asn Lys
Phe Gln Ala Ala Val Ala Leu Ala Leu 1 5 10
15 Leu Ser Arg Gly Ala Leu Gly Asp Ser Tyr Thr Asn Ser Thr Ser Ser 20
25 30 Ala Asp Leu Ser Ser Ile Thr Ser Val Ser Ser Ala Ser Ala Ser
Ala 35 40 45 Thr Ala Ser Asp Ser Leu Ser
Ser Ser Asp Gly Thr Val Tyr Leu Pro 50 55 60
Ser Thr Thr Ile Ser Gly Asp Leu Thr Val Thr Gly Lys Val Ile Ala 65
70 75 80 Thr Glu Ala Val Glu Val Ala Ala Gly Gly

Lys Leu Thr Leu Leu Asp 85 90 95
 Gly Glu Lys Tyr Val Phe Ser Ser Asp 100 105 <210> 2 <
 211> 430 <212> DNA <213> Saccharomyces cerevisiae <400> 2 gatcgtcata
 ttcactcttg ttctcataat agcagtcctaa gttttcatct ttgcaagctt 60 tactatttct
 ttctttttat tggtaaactc tcgcccatta caaaaaaaaa agagatgttc 120 aatcggttta
 acaaattcca agctgctgtc gctttggccc tactctctcg cggcgctctc 180 ggtgactctt
 acaccaatag cacctcctcc gcagacttga gttctatcac ttccgtctcg 240 tcagctagtg
 caagtgccac cgcttccgac tcactttctt ccagtgaagg taccgtttat 300 ttgccatcca
 caacaattag cggatgatctc acagttactg gtaaagtaat tgcaaccgag 360 gccgtggaag
 tcgctgccgg tggtaagttg actttacttg acggtgaaaa atacgtcttc 420 tcactctgatc
 430 <210> 3 <211> 117 <212> PRT <213> Saccharomyces cerevisiae <400>
 3 Met Thr Pro Tyr Ala Val Ala Ile Thr Val Ala Leu Leu Ile Val Thr 1 5
 10 15 Val Ser Ala Leu Gln Val Asn Asn Ser Cys Val Ala Phe Pro Pro
 Ser 20 25 30 Asn Leu Arg Gly Lys Asn
 Gly Asp Gly Thr Asn Glu Gln Tyr Ala Thr 35 40
 45 Ala Leu Leu Ser Ile Pro Trp Asn Gly Pro Pro Glu Ser Leu Arg Asp 50
 55 60 Ile Asn Leu Ile Glu Leu Glu Pro Gln Val Ala Leu Tyr Leu Leu
 Glu 65 70 75 80 Asn Tyr Ile Asn
 His Tyr Tyr Asn Thr Thr Arg Asp Asn Lys Cys Pro 85
 90 95 Asn Asn His Tyr Leu Met Gly Gly Gln Leu Gly Ser Ser Ser Asp
 Asn 100 105 110 Arg Ser Leu Asn Asp

```

115      <210>      4 <211>      424 <212>      DNA <213>      Saccharomyces cerevisiae <
400>      4 gatctcattg gattcaagag aaagaaactc tatactggcg ccaaattagc agtgtcaaat

60 ttcgaaaagg tgatgacgcc ctatgcagta gcaattaccg tggccttact aattgtaaca      120

gtgagcgcac tccagggtcaa caattcatgt gtcgcttttc cgccatcaaa tctcagaggc      180

aaaaatggag acggtactaa tgaacagtat gcaactgcac tactttctat tccctggaat      240

ggacctcctg agtcattgag ggatattaat cttattgaac tcgaaccgca agttgcactc      300

tatttgctcg aaaattatat taaccattac tacaacacca caagagacaa taagtgcctt      360

aataaccact acctaatggg agggcagttg ggtagctcat cggataatag gagtttgaac      420 gatc

424 <210>      5 <211>      104 <212>      PRT <213>      Saccharomyces cerevisiae <400>

5 Met Gln Phe Lys Asn Val Ala Leu Ala Ala Ser Val Ala Ala Leu Ser      5
9

14      19 Ala Thr Ala Ser Ala Glu Gly Tyr Thr Pro Gly Glu Pro Trp Ser

Thr      24      29      34 Leu Thr Pro Thr Gly Ser

Ile Ser Cys Gly Ala Ala Glu Tyr Thr Thr      39      44

49 Thr Phe Gly Ile Ala Val Gln Ala Ile Thr Ser Ser Lys Ala Lys Arg      54

59      64 Asp Val Ile Ser Gln Ile Gly Asp Gly Gln Val Gln Ala Thr Ser

Ala 69      74      79      84 Ala Thr Ala Gln

Ala Thr Asp Ser Gln Ala Gln Ala Thr Thr Thr Ala      89

94      99 Thr Pro Thr Ser Ser Glu Lys Ile      104

<210>      6 <211>      642 <212>      DNA <213>      Saccharomyces cerevisiae <400>      6

gatcccgctt agcccttcca gcttttcttt ttcccttttt gctacggctg agacacggtc      60

gcccaaaaga aacgggtcag cgtgtactgc gccaaaaaaa ttcgcgccga ttttaagctaa      120

```

```

acgtccacaa acaaaaacaa aaataagaaa taggttgaca gtgggtgaaa aattctcgaa      180
ggtttcatct ccaaacagtc agtatataag tattcgggaa agagagccaa tctatcttgt      240
ggtggtgcta tcttaacctt ctctttttgg cagtagtaat tgtaaataca gacacataaa      300
actatttcac tcgctaaact tacatctaaa atgcaattca aaaacgtcgc cctagctgcc      360
tccgttgctg ctctatccgc cactgcttct gctgaagggt acactccagg tgaaccatgg      420
tccaccttaa cccaaccggg ctccatctct tgtgggtgctg ccgaatacac taccaccttt      480
ggatttgctg ttcaagctat tacctcttca aaagctaaga gagacgttat ctctcaaatt      540
ggtgacggtc aagtccaagc cacttctgct gctactgctc aagccaccga tagtcaagcc      600
caagctacta ctaccgctac cccaaccagg tccgaaaaga tc                        642 <210>

7 <211>    50 <212>    PRT <213>    Hansenula polymorpha <400>    7 Met Arg Phe Ala
Glu Phe Leu Val Val Phe Ala Thr Leu Gly Gly Gly    5                                9
14                                19 Met Ala Ala Pro Val Glu Ser Leu Ala Gly Thr Gln Arg Tyr Leu
Val                                24                                29                                34 Gln Met Lys Glu Arg Phe
Thr Thr Glu Lys Leu Cys Ala Leu Asp Asp                                39                                44
49 Lys Ile    54 <210>    8 <211>    179 <212>    DNA <213>    Hansenula
polymorpha <400>    8 gatccgcttt ttattgcttt gctttgctaa tgagatttgc agaattcttg
gtggtatttg    60 ccacgttagg cgggggggatg gctgcaccgg ttgagtctct ggccgggacc
caacgttatc    120 tgggtgcaa at gaaggagcgg ttcaccacag agaagctgtg tgctttggac gacaagatc
179 <210>    9 <211>    71 <212>    PRT <213>    Saccharomyces cerevisiae <400>    9
Met Phe Asn Arg Phe Asn Lys Phe Gln Ala Ala Val Ala Leu Ala Leu    1                                5
10                                15 Leu Ser Arg Gly Ala Leu Gly Asp Ser Tyr Thr Asn Ser Thr Ser

```


Ser 20 25 30 Ala Asp Leu Ser Ser Ile
 Thr Ser Val Ser Ser Ala Ser Ala Ser Ala 35 40
 45 Thr Ala Ser Asp Ser Leu Ser Ser Ser Asp Gly Thr Val Tyr Leu Pro 50
 55 60 Ser Thr Thr Ile Ser Gly Asp 65 70 <210>
 10 <211> 329 <212> DNA <213> Saccharomyces cerevisiae <400> 10
 ggatccatgt tcaatcgttt taacaaattc caagctgctg tcgctttggc cctactctct 60
 cgcggcgctc tcggtgactc ttacaccaat agcacctcct ccgcagactt gagttctatc 120
 acttccgtct cgtcagctag tgcaagtgcc accgcttccg actcactttc ttccagtgac 180
 ggtaccgttt atttgccatc cacaacaatt agcgggtgac tcacagttac tggtaaagta 240
 attgcaaccg aggccgtgga agtcgctgcc ggtggtaagt tgactttact tgacggtgaa 300
 aaatacgtct tctcatctga tcctctaga 329 <210>
 11 <211> 35 <212> DNA <213> Artificial Sequence <220> <223> JH97 primer
 <400> 11 ccggccatta cggccgtgat gcacacaaga gtgag
 35 <210> 12 <211> 29 <212> DNA <213> Artificial Sequence <220> <223>
 JH119 primer <400> 12 ccggccgagg cggcctaagc ctaaggcag
 29 <210> 13 <211> 50 <212> DNA <213> Artificial Sequence <220> <223>
 JH99 primer <400> 13 gggcggccgc ctcggcccta gataaaaggt caatgacaaa cgaaactagc
 50 <210> 14 <211> 32 <212> DNA <213> Artificial Sequence <220> <223>
 JH100 primer <400> 14 ccgtcgactt actattttac ttcccttact tg
 32 <210> 15 <211> 37 <212> DNA <213> Artificial Sequence <220> <223>
 JH106 primer <400> 15 gcggccatta cggccgtgca cctacttcaa gttctac

37 <210> 16 <211> 37 <212> DNA <213> Artificial Sequence <220> <223>
 JH107 primer <400> 16 gcggccatta cggccgtgca cctacttcaa gttctac

37 <210> 17 <211> 26 <212> DNA <213> Artificial Sequence <220> <223>
 JH120 primer <400> 17 cgggatccgc acctacttca agttct

26 <210> 18 <211> 27 <212> DNA <213> Artificial Sequence <220> <223>
 JH121 primer <400> 18 cgggatcctg cacctacttc aagttct

27 <210> 19 <211> 28 <212> DNA <213> Artificial Sequence <220> <223>
 JH122 primer <400> 19 cgggatcctt gcacctactt caagttct

28 <210> 20 <211> 22 <212> DNA <213> Artificial Sequence <220> <223>
 JH123 primer <400> 20 ccattgaagg aaccaacaaa at

22 <210> 21 <211> 22 <212> DNA <213> Artificial Sequence <220> <223>
 JH124 primer <400> 21 attttgttgg ttccttcaat gg

22 <210> 22 <211> 29 <212> DNA <213> Artificial Sequence <220> <223>
 JH95 primer <400> 22 ggctcgagct attttacttc ccttacttg

29 <210> 23 <211> 23 <212> DNA <213> Artificial Sequence <220> <223>
 JH132 primer <400> 23 gggagctcat cgcttcgctg att

23 <210> 24 <211> 27 <212> DNA <213> Artificial Sequence <220> <223>
 JH137 primer <400> 24 ccgtcgactt aagttagtgt tgagatg

27 <210> 25 <211> 47 <212> DNA <213> Artificial Sequence <220> <223>
 HY22 primer <400> 25 gaacttgaag taggtgccct tttatctaga ggatcagatg agaagac

47 <210> 26 <211> 46 <212> DNA <213> Artificial Sequence <220> <223>

HY23 primer <400> 26 tcttctcatc tgatcctcta gataaaaggg cacctacttc aagttc
 46 <210> 27 <211> 46 <212> DNA <213> Artificial Sequence <220> <223>
 HY20 primer <400> 27 gaacttgaag taggtgccct tttatctaga ggatcgttca aactcc
 46 <210> 28 <211> 46 <212> DNA <213> Artificial Sequence <220> <223>
 HY21 primer <400> 28 ggagtttgaa cgatcctcta gataaaaggg cacctacttc aagttc
 46 <210> 29 <211> 47 <212> DNA <213> Artificial Sequence <220> <223>
 HY24 primer <400> 29 gaacttgaag taggtgccct tttatcaagg atcttgtcgt ccaaagc
 47 <210> 30 <211> 47 <212> DNA <213> Artificial Sequence <220> <223>
 HY25 primer <400> 30 gctttggacg acaagatcct tgataaaagg gcacctactt caagttc
 47 <210> 31 <211> 27 <212> DNA <213> Artificial Sequence <220> <223>
 JH143 primer <400> 31 cctctagaat caccgcta atgttgtg
 27 <210> 32 <211> 27 <212> DNA <213> Artificial Sequence <220> <223>
 JH142 primer <400> 32 cctctagagg tgctattgggt gtaagag
 27 <210> 33 <211> 27 <212> DNA <213> Artificial Sequence <220> <223>
 JH141 primer <400> 33 cctctagaac cgagagcgcc gcgagag
 27 <210> 34 <211> 25 <212> DNA <213> Artificial Sequence <220> <223>
 JH140 primer <400> 34 ggactagtct agataaaagg gcacc
 25 <210> 35 <211> 42 <212> DNA <213> Artificial Sequence <220> <223>
 HY38 primer <400> 35 gaatttttga aaattcaagg atccatgttc aatcgtttta ac
 42 <210> 36 <211> 33 <212> DNA <213> Artificial Sequence <220> <223>
 JH144 primer <400> 36 cctctagata aaaggacccc cctgggccct gcc

33 <210> 37 <211> 28 <212> DNA <213> Artificial Sequence <220> <223>

JH145 primer <400> 37 ggcagctgga tgtatatttac atggggag

28 <210> 38 <211> 43 <212> DNA <213> Artificial Sequence <220> <223>

HY17 primer <400> 38 gaacttgaag taggtgccct tttatcaagg atcttttcgg agc

43 <210> 39 <211> 43 <212> DNA <213> Artificial Sequence <220> <223>

HY18 primer <400> 39 gctccgaaaa gatccttgat aaaagggcac ctacttcaag ttc

43

